

# Localization of reinforced random walks

Pierre Tarrès<sup>1</sup>

January 18, 2013

## Abstract

We describe and analyze how reinforced random walks can eventually localize, i.e. only visit finitely many sites. After introducing vertex and edge self-interacting walks on a discrete graph in a general setting, and stating the main results and conjectures so far on the topic, we present martingale techniques that provide an alternative proof of the a.s. localization of vertex-reinforced random walks (VRRWs) on the integers on finitely many sites and, with positive probability, on five consecutive sites, initially proved by Pemantle and Volkov (1999,[11]).

Next we introduce the continuous time-lines representation (sometimes called Rubin construction) and its martingale counterpart, and explain how it has been used to prove localization of some reinforced walks on one attracting edge. Then we show how a modified version of this construction enables one to propose a new short proof of the a.s. localization of VRRWs on five sites on  $\mathbb{Z}$ .

*AMS 2000 subject classifications.* 60G50, 60J10, 60K35

*Key words and phrases.* reinforced walk, martingale, localization

*Running Head:* Localization of reinforced random walks

## 1 Introduction

Exploration of an environment, behaviour learning or cooperative interaction are instances of situations where the evolution depends on the whole history, either as a tendency to visit again “places” visited before or as a tendency to avoid them. In exploring an unknown city, streets that have been walked before may be considered as more attractive (safer, for instance) or repulsive (boring); learning the best choice

---

<sup>1</sup>CNRS, Université de Toulouse, Institut de Mathématiques, 118 route de Narbonne, 31062 Toulouse Cedex 9, France. On leave from the Mathematical Institute, University of Oxford. E-mail: [tarres@math.univ-toulouse.fr](mailto:tarres@math.univ-toulouse.fr)

among strategies giving random payoffs can be achieved by making random choices with an increasing preference towards the choices that pay more; and cooperation between micro-organisms, for instance, involves miming previously held behaviours. These situations naturally lend themselves to a modelization by self-interacting random processes.

The definition assumes we are given

- $(\Omega, \mathcal{F}, \mathbb{P})$  probability space,
- $(G; \sim)$  nonoriented locally finite graph,
- $(a_{i,j})_{i,j \in G, i \sim j}$  propensity matrix with positive entries, such that  $a_{i,j} > 0 \iff i \sim j$ ,
- $W : \mathbb{N}_0 \longrightarrow \mathbb{R}_+^*$  weight function.

The random process, called  $(X_n)_{n \in \mathbb{N}}$ , takes values in the set of vertices of  $G$ ; we let  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$  be the filtration of its past. For all  $v \in G$ ,  $n \in \mathbb{N} \cup \{\infty\}$ , let

$$Z_n(v) = \sum_{k=0}^n \mathbb{1}_{\{X_k=v\}} + 1 \quad (1)$$

be the number of visits to  $v$  up to time  $n$  plus one.

Then  $(X_n)_{n \in \mathbb{N}}$  is a Vertex Self-Interacting Random Walk (VSIRW) with starting point  $v_0 \in G$ , propensity matrix  $(a_{i,j})_{i \sim j}$  and weight function  $W$  if  $X_0 = v_0$  and, for all  $n \in \mathbb{N}$ , if  $X_n = i$  then

$$\mathbb{P}(X_{n+1} = j \mid \mathcal{F}_n) = \mathbb{1}_{i \sim j} \frac{a_{i,j} W(Z_n(j))}{\sum_{k \sim i} a_{i,k} W(Z_n(k))}. \quad (2)$$

An Edge Self-Interacting Random Walk (ESIRW) is defined similarly, replacing in (2) the numbers of visits to vertices  $l \sim i$  by those to the corresponding nonoriented edges  $\{i, l\}$ :

$$Z_n(\{i, l\}) := \sum_{k=1}^n (\mathbb{1}_{\{X_{k-1}=i, X_k=l\}} + \mathbb{1}_{\{X_{k-1}=l, X_k=i\}}) + 1. \quad (3)$$

We will define the Edge (resp. Vertex) Reinforced Random Walk as an ESIRW (resp. VSIRW) with linear  $W(n) = n + \Delta$ ,  $\Delta > -1$ : these processes were introduced by Coppersmith and Diaconis in 1986 [3].

In general, the asymptotic behaviour of self-interacting random walks greatly depends on the nature of the interaction. We will focus here on localization phenomena: the difficulty in their analysis lies in the fact that, before this localization occurs, the walk can concentrate on several disconnected clusters -separated by seldom visited sites- so that the relative numbers of visits follow a rather erratic dynamics, which is difficult to analyse.

For the study of strongly edge reinforced walks (i.e. ESIRW with reciprocally summable weight function  $W$ ), this technical difficulty can be partially overcome by

a simple argument, which allows to restrict the study to loop graphs (see Section 3). This argument cannot translate to vertex-reinforced random walks, which display localization on “richer” subsets, even on  $\mathbb{Z}$ , as we describe next.

## 1.1 Preliminary remarks

Let us start our study by the following simple preliminary results, which will enable us to gain more intuition on the behaviour of these walks: on one hand on the “simple” case of VSIRW on three vertices, and on the other hand on an easy but important property of ESIRWs.

We need to define the two following subsets  $R$  and  $R'$  of the graph, respectively called range and asymptotic range of the process  $(X_n)_{n \geq 0}$ :

$$\begin{aligned} R &:= \{v \in G \text{ s.t. } Z_\infty(v) \neq 1\} \\ R' &:= \{v \in G \text{ s.t. } Z_\infty(v) = \infty\}. \end{aligned}$$

We let  $\mathbf{Cst}(x_1, \dots, x_n)$  be a constant dependent only on  $x_1, \dots, x_n$ . The equalities and inclusions of probability events are understood to hold almost surely.

### 1.1.1 VSIRW on the three consecutive vertices $-1, 0$ and $1$

This walk is equivalent to the ESIRW on two non-oriented edges linking the same pair of vertices, which in turn can be seen as a  $W$ -urn process with two colours  $-1$  and  $1$ , defined as follows: we start with a certain number of balls of each colour (1 if  $X_0 = 0$ ) and, at each time step, we pick a ball of colour  $i \in \{-1, 1\}$  in the urn with a probability proportional to  $W(\text{number of balls of colour } i)$ , and put it back together with a ball of the same colour.

Assume  $X_0 = 0$  for simplicity. Let, for all  $n \in \mathbb{N}$ ,

$$\tilde{W}_\lambda(n) := \prod_{k=1}^{n-1} \left(1 + \frac{\lambda}{W(k)}\right), \quad (4)$$

$$W^*(n) := \sum_{k=1}^{n-1} \frac{1}{W(k)}, \quad \hat{W}(n) := \sum_{k=1}^{n-1} \frac{1}{W(k)^2}, \quad (5)$$

with the convention that  $W^*(1) = \hat{W}(1) := 0$  and  $\tilde{W}_\lambda(1) := 1$ .

For all  $\lambda > 0$ , let

$$A_n(\lambda) := \frac{\tilde{W}_{a_{0,-1}}(Z_n(1))}{\tilde{W}_{a_{0,1}}(Z_n(-1))},$$

and let

$$M_n := \left( \frac{d}{d\lambda} A_n(\lambda) \right)_{\lambda=0} = a_{0,-1} W^*(Z_n(1)) - a_{0,1} W^*(Z_n(-1)).$$

**Lemma 1.1** *The three processes  $(A_n(\lambda))_{n \geq 0}$ ,  $(A_n(\lambda)^{-1})_{n \geq 0}$  and  $(M_n)_{n \geq 0}$  are martingales.*

PROOF: Indeed, if  $X_n = 0$ , then

$$\begin{aligned} \frac{\mathbb{E}(A_{n+1}(\lambda) - A_n(\lambda) | \mathcal{F}_n)}{A_n(\lambda)} &= \frac{a_{0,1}W(Z_n(1))}{a_{0,1}W(Z_n(1)) + a_{0,-1}W(Z_n(-1))} \frac{a_{0,-1}}{W(Z_n(1))} \\ &\quad - \frac{a_{0,-1}W(Z_n(-1))}{a_{0,1}W(Z_n(1)) + a_{0,-1}W(Z_n(-1))} \frac{a_{0,1}}{W(Z_n(-1))} = 0. \end{aligned}$$

□

Let us now make use of Lemma 1.1 to analyze the asymptotic behaviour of the walk under the condition  $W(n) := (\Delta + n)^\rho$ ,  $\Delta > -1$ ,  $\rho \in \mathbb{R}$ :

1)  $\rho = 1$ , i.e.  $W(n) = n + \Delta$ .

Then  $W^*(n) = \sum_{k=1}^{n-1} 1/(k + \Delta)$ , and  $W^*(n) - \log n$  converges. On the other hand, for all  $n \geq 0$ ,

$$\sum_{k=0}^{n-1} (M_{k+1} - M_k)^2 = a_{0,-1}^2 \hat{W}(Z_k(1)) + a_{0,1}^2 \hat{W}(Z_k(-1)),$$

so that  $(M_n)_{n \geq 0}$  is bounded in  $L^2$ , hence converges a.s. and in  $L^2$  by Doob Lemma. Therefore  $a_{0,-1} \log(Z_n(1)) - a_{0,1} \log(Z_n(-1))$  converges, and

$$Z_n(1)^{a_{0,-1}} \underset{n \rightarrow \infty}{\sim} C Z_n(-1)^{a_{0,1}} \text{ a.s.}, \quad (6)$$

for a certain positive random variable  $C$ .

**a)** If  $a_{0,1} = a_{0,-1}$ , then the  $W$ -urn is a Pólya urn, and  $\frac{Z_n(1)}{Z_n(1) + Z_n(-1)}$  converges to a random variable  $\beta \in (0, 1)$ .

We can deduce from a classical result that  $\beta$  is a beta distribution with parameters  $(1 + \mathbb{I}_{\{X_0=1\}}\Delta, 1 + \mathbb{I}_{\{X_0=-1\}}\Delta)$  in general (but we assumed  $X_0 = 0$  here for simplicity).

**b)** If  $a_{0,1} \neq a_{0,-1}$ , note that the following urn process with the same colours  $-1$  and  $1$  could be analyzed by the technique above: at each time step, pick a ball in the urn with a probability proportional to the number of balls of that colour (as in Pólya urn), and we put it back together with, in conditional expectation,  $a_{0,1}$  (resp.  $a_{0,-1}$ ) if we picked colour  $1$  (resp.  $-1$ ).

The latter model is similar to Friedman urn [6], analyzed by Freedman in [5]. Note that we obtain the same martingale  $M_n$  (but not  $A_n(\lambda)$  for general  $\lambda$ ), and therefore the same asymptotics (6), if we assume for instance a bounded number of added balls.

2)  $\rho > 1$ , or more generally for any  $W$  satisfying  $\sum 1/W(n) < \infty$ .

Then the martingale  $(M_n)_{n \geq 0}$  converges a.s., as a difference of nondecreasing bounded sequences. On the other hand,

$$\{Z_\infty(1) = Z_\infty(-1) = \infty\} \subseteq \{M_\infty = 0\}.$$

It is possible to prove, using estimates of the variance of the increments, that  $\mathbb{P}(M_\infty = 0) = 0$  (see for instance [10]), so that only one of the two vertices 1 and  $-1$  is visited infinitely often almost surely. We will show that result by another technique in Section 3, in Proposition 3.1.

3)  $\rho < 1$  (not necessarily nonnegative).

Given  $\lambda > 0$ , using that  $(A_n(\lambda))_{n \geq 0}$  is a martingale of expectation 1, and that, for all  $x \geq 0$ ,  $0 \geq \log(1+x) - x \geq -x^2/2$ ,

$$\mathbb{E}[\exp(\lambda M_n)] \leq \exp(\lambda^2 \hat{W}(n)/2)$$

so that, by Chebychev inequality, for all  $a > 0$ ,

$$\mathbb{P}\left[M_n \geq a\sqrt{\hat{W}(n)}\right] \leq \exp(-\lambda a + \lambda^2 \hat{W}(n)/2) \leq \exp(-a^2/2),$$

choosing  $\lambda := a/\sqrt{\hat{W}(n)}$ .

Therefore, for all  $c > \sqrt{2}$ ,  $|M_n| \leq c\sqrt{\log n \hat{W}(n)}$  for large  $n$  a.s., so that

$$\frac{Z_n(1)}{Z_n(-1)} \xrightarrow{n \rightarrow \infty} \left(\frac{a_{0,1}}{a_{0,-1}}\right)^{(1-\rho)^{-1}}$$

Hence, on three vertices, the weakly reinforced walk ( $\rho \in (0, 1)$ ) behaves similarly as the self-repelling one ( $\rho < 0$ ), whereas the strongly reinforced walk -i.e. with  $W$  reciprocally summable- implies localization on two vertices.

In general, is strong reinforcement a necessary and sufficient condition for localization? The next subsection 1.1.2 and Section 1.3 will provide a partially positive answer for edge self-interaction (ESIRW), when  $W$  is nondecreasing. On the contrary, the results of Section 1.2 will highlight the dependence of the behaviour of VSIRWs on the graph structure, which also display localization on particular trapping patterns in the linear case  $W(n) = n + \Delta$ ,  $\Delta > -1$ .

### 1.1.2 ESIRW on $G$ connected, $W$ nondecreasing, $\sum_{n \in \mathbb{N}} \frac{1}{W(n)} = \infty$

**Proposition 1.1**  $\{|R'| \neq 0\} = \{R' = G\}$  a.s.

PROOF: Let  $t_n := t_n(x)$  be the  $n$ -th visit time to  $x$ , then

$$\sum_{z \sim x} Z_{t_n(x)}(\{x, z\}) = 2n + a,$$

where  $a := \mathbb{I}_{\{X_0 \neq x\}} - 2 + |\{z \in G : z \sim x\}|$ . Hence, for all  $z \sim x$  and  $n \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{P}(X_{t_n+1} = z | \mathcal{F}_{t_n}) \mathbb{I}_{\{t_n < \infty\}} &\geq \frac{a_{x,z}}{\sum_{y \sim x} a_{x,y}} \frac{W(0)}{W(2n+a)} \\ &\geq \mathbb{I}_{\{t_n < \infty\}} W(0) \text{Cst}(a_{x,y})_{y \sim x} \left( \frac{1}{W(2n+a)} + \frac{1}{W(2n+1+a)} \right), \end{aligned}$$

using that  $W$  is nondecreasing.

Therefore, using conditional Borel-Cantelli Lemma 5.1,

$$\begin{aligned} \{Z_\infty(x) = \infty\} &\subseteq \left\{ \sum_{n \in \mathbb{N}} \mathbb{P}(X_{t_n+1} = z | \mathcal{F}_{t_n}) \mathbb{I}_{\{t_n < \infty\}} = \infty \right\} \\ &= \left\{ \sum_{n \in \mathbb{N}} \mathbb{I}_{\{X_{t_n+1}=z\}} \mathbb{I}_{\{t_n < \infty\}} = \infty \right\} \subseteq \{Z_\infty(z) = \infty\} \text{ a.s.} \end{aligned}$$

□

**Remark 1.1** The condition that  $W$  should be nondecreasing is important in Proposition 1.1. The following counterexample was proposed by Sellke [12]: if  $\sum \frac{1}{W(2k)} = \infty$  and  $\sum \frac{1}{W(2k+1)} < \infty$ ,  $a_{i,j} = \mathbb{I}_{\{i \sim j\}}$ ,  $G = \mathbb{R}^d$  and  $X_0 = 0$ , then

$$\mathbb{P}(\forall n \in \mathbb{N}, X_{2n} = 0) > 0.$$

This result will be a direct consequence of the time-lines construction in Section 3.

## 1.2 Localization results, VSRIW

### 1.2.1 $G = \mathbb{Z}$ , $a_{i,j} = \mathbb{I}_{i \sim j}$ , $W(n) = \Delta + n$

Pemantle and Volkov [11] showed that the walk a.s. visits only finitely many vertices, and that it localizes on any set of five consecutive sites with positive probability. Tarrès [13] proved that localization on five sites is in fact the a.s. behaviour.

**Theorem 1 (Pemantle and Volkov, [11])**  $|R'| < \infty$  a.s. and, for any  $x \in \mathbb{Z}$ ,  $\mathbb{P}(R' = \{x-2, x-1, x, x+1, x+2\}) > 0$ .

**Theorem 2 (Tarrès, [13])**  $|R'| = 5$  a.s.

We propose here some new proofs of these results: Theorem 1 in Section 2 (see Propositions 2.1 and 2.2), using techniques derived from [13], and Theorem 2 in Section 4, partially adapting a continuous-time equivalent of the random walk, originally introduced in [4, 12].

**1.2.2**  $G = \mathbb{Z}$ ,  $a_{i,j} = \mathbb{1}_{i \sim j}$ ,  $W(n) \sim n^\alpha$

**Theorem 3 (Volkov, [14])** *Suppose that  $0 < \lim_{k \rightarrow \infty} w_k/k^\alpha < \infty$ . Then*

(a) *If  $\alpha < 1$ , then  $R' \in \{0, \infty\}$*

(b) *If  $\alpha > 1$ , then  $|R'| = 2$ .*

**1.2.3 General  $(G, \sim)$  locally finite,  $(a_{i,j})_{i \sim j}$  symmetric and  $W(n) = \Delta + n$**

Under a symmetric propensity matrix, the vertex-reinforced random walk localizes with positive probability on a class of complete  $d$ -partite subgraphs with possible loops plus their outer boundary. We need to introduce some notation, in order to describe further these trapping subsets.

Given a subset  $R$  of  $G$ , we let

$$\partial R = \{j \in G \setminus R : j \sim R\}$$

be the *outer boundary* of  $R$ .

For any  $x = (x_i)_{i \in G} \in \mathbb{R}^G$ , let

$$S(x) := \{i \in G / x_i \neq 0\}$$

be its support. Let

$$\Delta := \left\{ x \in \mathbb{R}_+^G \text{ s.t. } |S(x)| < \infty \text{ and } \sum_{i \in G} x_i = 1 \right\}$$

be the nonnegative simplex restricted to elements  $x$  of finite support.

For all  $x \in \Delta$ , let

$$N_i(x) := \sum_{j \in G, j \sim i} a_{i,j} x_j, \quad H(x) = \sum_{i,j \in G, i \sim j} a_{i,j} x_i x_j = \sum_{i \in G} x_i N_i(x). \quad (7)$$

For all  $n \in \mathbb{N}$ , let

$$x_n = \left( \frac{Z_n(i) - 1}{n} \right)_{i \in G}$$

be the vector of density of occupation of the random walk at time  $n$ , which has finite support and takes values in  $\Delta$ .

The following definition introduces “good candidates” for the limiting density of occupation of the random walk.

**Definition 1.1** *For all  $x \in \Delta$ , let  $(\mathbf{P})_x$  be the following predicate:*

$$\max \left( \text{Sp} [a_{i,j} - 2H(x)]_{i,j \in S(x)} \right) \leq 0, \quad \max \{N_i(x) - H(x), i \in \partial S(x)\} < 0.$$

**Theorem 4 (Benaïm and Tarrès, [1])** *Given  $x \in \Delta$  such that  $(\mathbf{P})_x$  holds and for any neighbourhood  $\mathcal{N}(x)$  of  $x$  in  $\Delta$ , there is with positive probability  $y \in \mathcal{N}(x)$  with  $S(y) = S(x)$ , such that the following three events occur:*

- (i)  $R' = S(x) \cup \partial S(x)$
- (ii)  $x_n \rightarrow y$
- (iii)  $\forall i \in \partial S(x), Z_n(i)/n^{N_i(x)/H(x)} \rightarrow C_i \in (0, \infty)$  (random).

Assumption  $(\mathbf{P})_x$  in Definition 7 describes stable equilibria of the ordinary differential equation

$$\frac{dx}{dt} = F(x), \quad (8)$$

where

$$F(x) = (x_i[N_i(x) - H(x)])_{i \in G}. \quad (9)$$

also known as the linear *replicator* equation in population genetics and game theory. Up to an adequate rescaling in time, we can indeed show that  $(x_k)_{k \in \mathbb{N}}$  approximate of this ODE under certain assumptions.

The support of these equilibria satisfies some properties, described in the following Lemma 1.2. In the context of population dynamics, they inform on the structure of the surviving species, depending on the nature of the graph.

**Definition 1.2** *Given  $d \geq 1$ , subgraph  $(S, \sim)$  of  $(G, \sim)$  will be called a complete  $d$ -partite graph with possible loops, if  $(S, \sim)$  is a  $d$ -partite graph on which some loops have possibly been added. That is*

$$S = V_1 \cup \dots \cup V_d$$

with

- (i)  $\forall p \in \{1, \dots, d\}, \forall i, j \in V_p, \text{ if } i \neq j \text{ then } i \not\sim j.$
- (ii)  $\forall p, q \in \{1, \dots, d\}, p \neq q, \forall i \in V_p, \forall j \in V_q, i \sim j.$

**Definition 1.3** *For all  $S \subseteq G$ , let  $(\mathbf{P})_S$  be the following predicate:*

- $(\mathbf{P})_S(\mathbf{a})$   $(S, \sim)$  is a complete  $d$ -partite graph with possible loops.
- $(\mathbf{P})_S(\mathbf{b})$  If  $i \sim i$  for some  $i \in S$ , then the partition containing  $i$  is a singleton.
- $(\mathbf{P})_S(\mathbf{c})$  If  $V_p, 1 \leq p \leq d$  are its  $d$  partitions, then for all  $p, q \in \{1, \dots, d\}$  and  $i, i' \in V_p, j, j' \in V_q, a_{i,j} = a_{i',j'}$ .

**Lemma 1.2** (Benaïm and Tarrès, [1]) *For all  $x \in \Delta$ ,  $(\mathbf{P})_x$  implies  $(\mathbf{P})_{S(x)}$ .*



### 1.3 Localization results, ESIRW

Let (H) be the following condition on  $W$ :

$$\sum_{k \in \mathbb{N}} \frac{1}{W(k)} < \infty.$$

We know from Proposition 1.1 that, if (H) does not hold and  $W$  is nondecreasing, then the walk is either transient or recurrent on all vertices, assuming the graph  $G$  is connected.

On the other hand, it is easy to show that (H) implies localization on a single edge with positive probability. Sellke [12] conjectured in 1994 that this should occur with probability one on any graph of bounded degree, and proved the statement on graphs without odd cycles. Then Limic and Tarrès [9] showed in 2007 that this conjecture indeed holds if  $W$  is nondecreasing (Limic [8] solved the case  $W(k) = (k+1)^\rho$  in 2003).

**Theorem 5 (Sellke [12], Theorem 3, Limic [8], Lemmas 1-2, Corollaries 1-2)** *If  $(G, \sim)$  has bounded degree and contains no odd cycles, then (H) implies  $|R'| = 2$  a.s.*

**Theorem 6 (Limic and Tarrès [9], Corollary 3)** *If  $(G, \sim)$  has bounded degree and  $W$  is nondecreasing, then (H) implies  $|R'| = 2$  a.s.*

We explain the key ingredient for the proof of Theorem 5 in Section 3.2.

## 2 VRRW on $\mathbb{Z}$ , first localization results

The goal of this section is to prove the following Propositions 2.1 and 2.2, which will in particular imply Theorem 1.

**Proposition 2.1** *(Pemantle and Volkov, 1999, [11])*

- (a) *For all  $x \in \mathbb{Z}$ ,  $\alpha \in (0, 1)$ ,  $\epsilon > 0$ ,*  

$$\mathbb{P}(\{R' = \{x-2, x-1, x, x+1, x+2\}\} \cap \{\alpha_\infty^-(x) \in (\alpha - \epsilon, \alpha + \epsilon)\}) > 0.$$
- (b)  *$|R| < \infty$  a.s.*

**Proposition 2.2**

- (a) *(Bienvenüe, 1999, [2])*  

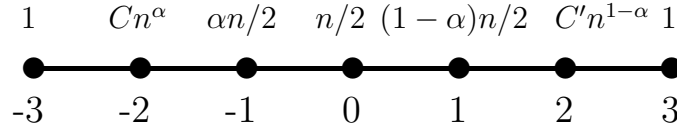
$$\{R' = \{x-2, x-1, x, x+1, x+2\}\}$$

$$\subseteq \{\exists \alpha_\infty^-(x) := \lim \alpha_n^-(x) \in (0, 1)\} \cap \{\log Z_n(x \pm 2) / \log Z_n(x) \xrightarrow[n \rightarrow \infty]{} \alpha_\infty^\pm(x)\} \text{ a.s.}$$
- (b) *(Pemantle and Volkov, 1999, [11])  $|R'| \geq 5$  a.s.*

We will assume that  $\Delta := 0$  in the remainder of this survey, so that  $W(n) = n$ ; the proofs obviously carry on to the general case, by replacing  $Z(\cdot)$  by  $Z(\cdot) + \Delta$ , and by defining the following function  $h$  accordingly, on  $\Delta + \mathbb{N}$  instead of  $\mathbb{N}$ . Also recall that  $a_{i,j} = \mathbb{I}_{i \sim j}$  here.

Let us first explain the heuristics of the localization on five vertices with positive probability. Let us assume for instance that the walk started at 0, and that we are in the following configuration at time  $n$ , represented by the figure below: site 0 has been visited  $n/2$  times, its neighbours  $-1$  and  $1$  have shared the other half of time, and have been visited respectively roughly  $\alpha n/2$  and  $(1 - \alpha)n/2$  times for some constant  $\alpha \in (0, 1)$ ; that the two sites  $-2$  and  $2$  have been visited of the order of  $Cn^\alpha$  and  $C'n^{1-\alpha}$  and, finally that  $-3$  and  $3$  have not been visited yet.

The numbers above the sites on the figure represent the estimates of the numbers of visits (plus one).



Assume  $X_n = -1$  for instance. First, the configuration would not be viable asymptotically for an edge-reinforced random walk (ERRW), since

$$\mathbb{P}(X_{n+2} = -3 | \mathcal{F}_n) \underset{n \rightarrow \infty}{\sim} \frac{2Cn^\alpha}{\alpha n} \frac{1}{2Cn^\alpha} \underset{n \rightarrow \infty}{\sim} \frac{1}{\alpha n}$$

so that  $-3$  would eventually be visited, by conditional Borel-Cantelli lemma.

On the contrary, for the vertex-reinforced random walk we are considering, the same computation yields

$$\mathbb{P}(X_{n+2} = -3 | \mathcal{F}_n) \underset{n \rightarrow \infty}{\sim} \frac{Cn^\alpha}{n/2} \frac{1}{\alpha n/2} \underset{n \rightarrow \infty}{\sim} \frac{\text{Cst}(C, \alpha)}{n^{2-\alpha}}.$$

Therefore, with lower bounded probability, the sites  $-3$  and  $3$  will never be visited, as long as the same asymptotics holds for the visits to the other sites. Now, under these assumptions:

- The visits to  $-2$  and  $2$  are seldom, so that the respective visits to  $-1$  and  $1$  can almost be estimated by considering the walk on the three vertices  $-1, 0$  and  $1$ , restricted to its moves to and from  $0$ , described in Section 1.1.1: we are in Case 1)a), so that  $Z_n(-1)/(Z_n(-1) + Z_n(1))$  should remain close to  $\alpha$ .
- The respective visits to  $-2$  (similarly  $2$ ) and  $0$  can be seen as stemming from a Friedman urn model, Case 2)a) of Section 1.1.1. Indeed, starting from  $-1$ , the sites  $-2$  and  $0$  are chosen proportionally to their numbers of visits. But, if  $-2$  is picked then the walk immediately comes back to  $-1$ , whereas if  $0$  is chosen, the expectation of the number visits before coming back to  $-1$  is roughly  $\alpha^{-1}$ : (6) would provide convergence of  $Z_n(-2)/Z_n(0)^\alpha$ .

The following results indeed justify this heuristics, and will also be useful in the proof of a.s. localization on five points. They are similar to those developped in Sections 1, 2 and 3 of [13].

Let, for all  $n \in \mathbb{N}$ ,

$$h(n) := \sum_{k=1}^{n-1} \frac{1}{k},$$

with the convention that  $h(1) := 0$ . For all  $n \in \mathbb{N}_0$  and  $x \in \mathbb{Z}$ , denote

$$\begin{aligned} Z_n^\pm(x) &:= \sum_{k=1}^n \mathbb{I}_{\{X_{k-1}=x, X_k=x\pm 1\}}, \\ \alpha_n^\pm(x) &:= \frac{Z_n(x \pm 1)}{Z_n(x-1) + Z_n(x+1)}, \\ Y_n^\pm(x) &:= \sum_{k=1}^n \mathbb{I}_{\{X_{k-1}=x, X_k=x\pm 1\}} \frac{1}{Z_{k-1}(x \pm 1)}, \end{aligned}$$

also

$$\begin{aligned} Y_n(x) &:= \sum_{k=1}^n \mathbb{I}_{\{X_{k-1}=x\}} \frac{1}{Z_{k-1}(x-1) + Z_{k-1}(x+1)}, \\ \hat{Y}_n^\pm(x) &:= Y_n^\pm(x) - Y_n(x), \end{aligned}$$

which are respectively the previsible and martingale part in the Doob decomposition of  $Y_n^\pm(x)$ , and finally

$$Y_\infty^\pm(x) := \lim_{n \rightarrow \infty} Y_n^\pm(x), \quad Y_\infty(x) := \lim_{n \rightarrow \infty} Y_n(x).$$

Given  $(a_n), (b_n)$  random processes on  $\mathbb{R}$ , we write  $a_n \equiv b_n$  iff  $a_n - b_n$  converges a.s. Let us define the probability event

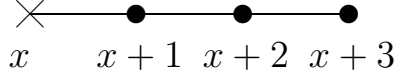
$$\Upsilon(x) := \{Y_\infty(x) < \infty\}$$

and, for any finite sequence  $(x_i)_{1 \leq i \leq n}$  taking values in  $\mathbb{Z}$ , the event

$$\Upsilon((x_i)_{1 \leq i \leq n}) = \bigcap_{1 \leq i \leq n} \Upsilon(x_i).$$

## 2.1 “Pólya urn” estimates

The event  $\Upsilon(x)$  corresponds to the event that  $x$  is “seldom” visited (represented by a cross on the figure), hence “neutral” with respect to its neighbours, in the following sense: the respective visits to  $x+1$  and  $x+3$  starting from  $x+2$  can be seen be seen similar to those of a Pólya urn model (see Section 1.1.1, study of case 1)a)), perturbed by the visits from  $x$  and  $x+4$ :  $\Upsilon(x)$  implies that the visits from  $x$  do not act upon the asymptotic behaviour of  $\alpha_n^-(x+2)$ , as stated in Corollary 2.1.



**Proposition 2.3** *For all  $x \in \mathbb{Z}$ ,*

- (a)  $\hat{Y}_n^\pm(x) = Y_n^\pm(x) - Y_n(x)$  *is a martingale, converging a.s. and in  $L^2$*
- (b)  $Y_n^\pm(x) \equiv Y_n(x)$
- (c)  $\mathbb{E}((\hat{Y}_n(x) - \hat{Y}_\infty(x))^2 | \mathcal{F}_n) \leq \text{Cst } Z_n(x \pm 1)^{-1}$
- (d)  $Y_n^+(x-1) + Y_n^-(x+1) = h(Z_n(x)) - h(1 + \mathbb{I}_{\{X_0=x\}}) \equiv \log Z_n(x)$

PROOF: It follows from its definition that  $(\hat{Y}_n^\pm(x))_{n \geq 0}$  is a martingale. Now

$$\begin{aligned} \text{Var}(\hat{Y}_{n+1}^\pm(x) | \mathcal{F}_n) &= \text{Var}(Y_{n+1}^\pm(x) | \mathcal{F}_n) \leq \mathbb{E}((Y_{n+1}^\pm(x) - Y_n^\pm(x))^2 | \mathcal{F}_n) \\ &= \mathbb{E}\left(\frac{\mathbb{I}_{\{X_n=x, X_{n+1}=x \pm 1\}}}{Z_n(x \pm 1)^2} | \mathcal{F}_n\right) \end{aligned}$$

Hence, for all  $m \geq n$ ,

$$\mathbb{E}((Y_n^\pm(x) - Y_m^\pm(x))^2 | \mathcal{F}_n) \leq \mathbb{E}\left(\sum_{k=n}^{\infty} \frac{\mathbb{I}_{\{X_k=x, X_{k+1}=x \pm 1\}}}{Z_k(x \pm 1)^2} | \mathcal{F}_n\right) \leq \sum_{l=Z_n(x \pm 1)}^{\infty} \frac{1}{l^2} \leq \frac{\text{Cst}}{Z_n(x \pm 1)}.$$

This implies (a)-(c); (d) follows from definitions.  $\square$

Let, for all  $t \in (0, 1)$ ,

$$f(t) := \log\left(\frac{t}{1-t}\right).$$

Let, for all  $a > 0$ ,  $m \geq 0$ ,  $y \in \mathbb{Z}$ , let  $\mathcal{E}_{m,a}(y)$  be the event

$$\mathcal{E}_{m,a}(y) := \left\{ \sup_{m \leq k \leq n < \infty} [f(\alpha_n^-(y)) - f(\alpha_k^-(y)) - (Y_n(y-2) - Y_k(y-2))] \geq a \right\}. \quad (10)$$

**Corollary 2.1** *For all  $x \in \mathbb{Z}$ ,  $n \geq m$ ,  $a \geq \text{Cst}$ ,*

- (a)  $\Upsilon(x) \subseteq \{\exists \alpha_\infty^-(x+2) := \lim \alpha_n^-(x+2) \in [0, 1)\} \cap \{\log f(\alpha_n^-(x+2)) \equiv -Y_n^-(x+4)\}$
- (b)  $\Upsilon(x) \cap \{\alpha_\infty^-(x+2) > 0\} \subseteq \Upsilon(x+4)$
- (c)  $\mathbb{P}(\mathcal{E}_{m,a}(x+2) | \mathcal{F}_m) \leq \text{Cst}(a) (Z_m(x+1)^{-1} + Z_m(x+3)^{-1})$
- (d)  $\Upsilon(x-1, x+1) \subseteq \{Z_\infty(x) < \infty\}$

PROOF: **(a)-(b)** By Proposition 2.3 **(b)** and **(d)**, a.s. on  $\Upsilon(x) = \{Y_\infty^-(x) < \infty\}$ ,

$$\log Z_n(x+1) \equiv Y_n^+(x) + Y_n^-(x+2) \equiv Y_n^+(x+2) \equiv \log Z_n(x+3) - Y_n^-(x+4),$$

so that

$$\log f(\alpha_n^-(x+2)) = \log \frac{Z_n(x+1)}{Z_n(x+3)} \equiv -Y_n^-(x+4).$$

**(c)** By Proposition 2.3 **(d)**,

$$\begin{aligned} h(Z_n(x+1)) - h(1 + \mathbb{I}_{\{X_0=x+1\}}) &= Y_n^+(x) + Y_n^-(x+2) \\ &= Y_n(x) + Y_n^+(x+2) + \hat{Y}_n^+(x) + \hat{Y}_n^-(x+2) - \hat{Y}_n^+(x+2) \\ &= Y_n(x) + h(Z_n(x+3)) - h(1 + \mathbb{I}_{\{X_0=x+3\}}) - Y_n^-(x+4) \\ &\quad + \hat{Y}_n^+(x) + \hat{Y}_n^-(x+2) - \hat{Y}_n^-(x+2). \end{aligned} \quad (11)$$

In order to estimate the probability of  $\mathcal{E}_{m,a}(x+2)$  we observe that, if  $a \geq \text{Cst}$ , then

$$\begin{aligned} f(\alpha_n^-(x+2)) - f(\alpha_m^-(x+2)) &= \log \left( \frac{Z_n(x+1)}{Z_m(x+1)} \right) - \log \left( \frac{Z_n(x+3)}{Z_m(x+3)} \right) \\ &\leq h(Z_n(x+1)) - h(Z_m(x+1)) - [h(Z_n(x+3)) - h(Z_m(x+3))] + \frac{a}{4}, \end{aligned}$$

using that, for any large  $q \geq p$ ,

$$-\frac{\text{Cst}}{p} \leq \log \frac{q}{p} - (h(q) - h(p)) \leq 0, \quad (12)$$

as  $1/p - 1/p^2 \leq \log((p+1)/p) \leq 1/p$  for large  $p$ .

We subtract identities (11) at times  $n$  and  $m$ , and note that the terms  $\hat{Y}_n^+(x) - \hat{Y}_m^+(x)$ ,  $\hat{Y}_n^-(x+2) - \hat{Y}_m^-(x+2)$  and  $\hat{Y}_m(x+2) - \hat{Y}_n(x+2)$  can be upper bounded by  $a/4$  with large probability, by Doob and Chebychev inequalities, and Proposition 2.3 **(c)**, for instance:

$$\mathbb{P} \left( \sup_{n \geq m} (\hat{Y}_n^+(x) - \hat{Y}_m^+(x)) \geq \frac{a}{4} \middle| \mathcal{F}_m \right) \leq \frac{\text{Cst}}{a^2} \mathbb{E}((\hat{Y}_n(x) - \hat{Y}_\infty(x))^2 | \mathcal{F}_n) \leq \frac{\text{Cst}}{a^2} \frac{1}{Z_m(x \pm 1)},$$

which completes the proof.

**(d)** It is a direct consequence of Proposition 2.3 **(b)** and **(d)**: a.s. on  $\Upsilon(x-1, x+1)$ ,

$$\log Z_\infty(x) = Y_\infty^+(x-1) + Y_\infty^-(x+1) < \infty.$$

□

## 2.2 “Friedman urn” estimates

The main goal of this subsection is to provide estimates for the number of visits to  $x + 1$  when  $\alpha_n^-(x + 3)$  converges to  $\alpha_\infty^-(x + 3) > 0$ , and  $\Upsilon(x - 1)$  holds, i.e.  $x - 1$  is “seldom” visited. The Friedman urn dynamics, which arises for sites  $x$  and  $x + 2$ , appears through the following calculation, justified rigourously in this section:

$$\begin{array}{ccccccc} \times & & \times & & \bullet & & \bullet & & \bullet \\ x-1 & & x & & x+1 & & x+2 & & x+2 \end{array}$$

$$\begin{aligned} \log Z_n(x) &\equiv Y_n^-(x+1) \equiv Y_n^+(x+1) = \sum_{k=1}^n \frac{\mathbb{I}_{\{X_{k-1}=x+1, X_k=x+2\}}}{Z_{k-1}(x+2)} \\ &\approx \sum_{k=1}^n \frac{\mathbb{I}_{\{X_{k-1}=x+2, X_k=x+1\}}}{Z_{k-1}(x+2)} \approx \sum_{k=1}^n \frac{\mathbb{I}_{\{X_k=x+2\}}}{Z_{k-1}(x+2)} \alpha_k^-(x+2) \approx \alpha_\infty^-(x+2) \log Z_n(x+2). \end{aligned}$$

For all  $x \in \mathbb{Z}$ ,  $n \in \mathbb{N}$ ,  $\alpha > 0$ , let

$$\begin{aligned} U_{n,\pm}^+(x) &:= \sum_{k=1}^n \frac{\mathbb{I}_{\{X_{k-1}=x, X_k=x\pm 1\}}}{Z_{k-1}(x)} \\ U_{n,\pm}^-(x) &:= \sum_{k=1}^n \frac{\mathbb{I}_{\{X_{k-1}=x, X_k=x\mp 1\}}}{Z_{k-1}(x)} \frac{Z_{k-1}(x+1)}{Z_{k-1}(x-1)}, \end{aligned}$$

and let

$$\begin{aligned} U_{n,\pm}(x) &:= \sum_{k=1}^n \frac{\mathbb{I}_{\{X_{k-1}=x\}}}{Z_{k-1}(x)} \alpha_{k-1}^\pm(x) \\ \hat{U}_{n,\pm}^\pm(x) &:= U_{n,\pm}^\pm(x) - U_{n,\pm}(x) \end{aligned}$$

be respectively the martingale and previsible parts of the Doob decomposition of  $U_{n,\pm}^\pm(x)$ .

Note that the processes  $U_{n,\pm}^-(x)$  and  $\hat{U}_{n,\pm}^-(x)$  will not be used in this section; they will be necessary tools to the proof of a.s. localization in Section 4.

**Proposition 2.4** For all  $x \in \mathbb{Z}$ ,  $C, a, \gamma > 0$ ,  $n \in \mathbb{N}$ ,

- (a)  $n \mapsto Y_n^\pm(x) + \frac{\mathbb{I}_{\{\pm X_n \leq \pm x\}}}{Z_{n-1}(x \pm 1)} - U_{n,\mp}^+(x \pm 1)$  is a.s. constant
- (b)  $\hat{U}_{n,\pm}^+(x)$  converges a.s. and in  $L^2$ , and
$$\mathbb{E}((\hat{U}_{n,\pm}^+(x) - \hat{U}_{\infty,\pm}^+(x))^2 | \mathcal{F}_n) \leq \text{Cst } Z_n(x)^{-1}$$
- (c)  $\Upsilon(x-1) \cap \{Z_\infty(x) = \infty\} \cap \{\limsup \alpha_n^-(x+2) \leq \gamma\}$   
 $\subseteq \Upsilon(x-1, x) \cap \{\log Z_n(x) \underset{n \rightarrow \infty}{\sim} \alpha_\infty^-(x+2) \log Z_n(x+2)\}$
- (d) Assume  $Z_m \leq CZ_m(x+2)^\alpha$ , and let  
 $T := \inf\{n \geq m \text{ s.t. } \alpha_n^-(x+2) \geq \gamma \text{ or } X_n = x-1\}$ . Then  

$$\mathbb{P}\left(\sup_{T > n \geq m} \frac{Z_n(x)}{Z_n(x+2)^\gamma} \geq e^a C | \mathcal{F}_m\right) \leq \text{Cst}(a)(Z_m(x)^{-1} + Z_m(x+2)^{-1})$$

PROOF: (a) Assume  $\pm = +$  for simplicity. Now the function of  $n$  only changes when  
-  $X_{n-1} = x$  and  $X_n = x+1$ , in which case  $Y_n^+(x)$  increases by  $Z_{n-1}(x+1)^{-1}$   
-  $X_{n-1} = x+1$  and  $X_n = x$ , in which case  $U_{n,-}^+(x+1)$  increases by  $Z_{n-1}(x+1)^{-1}$   
(b) is similar to Proposition 2.3 (c)  
(c) Note that

$$\begin{aligned} \log Z_n(x) &\equiv Y_n^+(x-1) + Y_n^-(x+1) \equiv Y_n^+(x-1) \equiv U_{n,-}^+(x+2) \equiv U_{n,-}(x+2) \\ &\equiv \sum_{k=1}^n \frac{\mathbb{I}_{\{X_{k-1}=x+2\}}}{Z_{k-1}(x+2)} \leq (\gamma + \epsilon) \log Z_n(x+2) \end{aligned}$$

for large  $n$  if  $\limsup \alpha_n^-(x+2) \leq \gamma$  and  $Z_\infty(x) = \infty$ . This implies  $\alpha_k^-(x+1) \leq (Z_n(x) + Z_n(x+2))^{\gamma+\epsilon-1} \leq Z_n(x+1)^{\gamma+\epsilon-1}$  for large  $n \in \mathbb{N}$ , so that

$$Y_\infty(x) \equiv Y_\infty^+(x) \equiv U_{n,-}(x+1) = \sum_{k=1}^n \frac{\mathbb{I}_{\{X_{k-1}=x+1\}}}{Z_{k-1}(x+1)} \alpha_{k-1}^-(x+1) < \infty.$$

(d) Similarly as in Corollary 2.1 (c), outside of an event of probability less than  $\text{Cst}(a)(Z_m(x)^{-1} + Z_m(x+2)^{-1})$ , we have

$$\log \left( \frac{Z_n(x)}{Z_m(x)} \right) \leq Y_n^+(x-1) - Y_m^+(x-1) + a + \sup_{m \leq k \leq n} \alpha_{k-1}^\pm(x+2) \log \left( \frac{Z_n(x+2)}{Z_m(x+2)} \right).$$

□

## 2.3 Conclusions

### 2.3.1 Proof of Proposition 2.2

(a) A.s. on  $\{Z_\infty(x+3) < \infty\} \cap \{Z_\infty(x-3) < \infty\}$ ,  $\Upsilon(x-3, x-2, x+2, x+3)$  holds by Proposition 2.3 (b). Hence  $\alpha_n^-(x) \xrightarrow{n \rightarrow \infty} \alpha_\infty^-(x) \in [0, 1)$  by Corollary 2.1 (a),

which also implies  $\alpha_\infty^-(x) \in [0, 1)$  by considering the sites in the reverse order, so that  $\alpha_\infty^-(x) \in (0, 1)$ . Finally Proposition 2.4 (c) completes the proof.

(b) A.s.  $\{R' \subseteq \{x, x+1, x+2, x+3\}\} \cap \{Z_\infty(x) = \infty\}$ , Corollary 2.1 (a) implies  $\alpha_n^-(x+2) \xrightarrow{n \rightarrow \infty} \alpha_\infty^-(x+2) \in [0, 1)$ ; subsequently,  $\alpha_n^-(x+1) \xrightarrow{n \rightarrow \infty} 0$  by Proposition 2.4 (c). On the other hand,  $\Upsilon(x+3)$  holds, so that  $\alpha_\infty^+(x+1) \xrightarrow{n \rightarrow \infty} \alpha_\infty^+(x+1) \in [0, 1)$ , which is contradictory.

### 2.3.2 Proof of localization with positive probability

We will prove the following result of localization on a half-axis.

**Corollary 2.2** *For all  $x \in \mathbb{Z}$ ,  $\gamma \in [0, 1)$ ,  $\epsilon > 0$ ,  $C > 0$ ,  $m \in \mathbb{N}$ , assume  $Z_m(x \pm 3) = 1$ ,  $Z_m(x \pm 2) \leq CZ_m(x)^\gamma$  and  $\alpha_m^\pm(x) \leq \gamma$ . Then*

$$\mathbb{P} \left( \{Z_\infty(x \pm 3) \neq 1\} \cup \left\{ \sup_{n \geq m} \alpha_n^\pm(x) > \gamma + \epsilon \right\} \middle| \mathcal{F}_m \right) \leq \text{Cst}(\gamma, \epsilon, C) (Z_m(x)^{-\gamma} + Z_m(x \pm 1)^{\gamma + \epsilon - 1}).$$

PROOF: Assume  $\pm := -$ ,  $X_m \geq x - 1$  and  $\epsilon < 1 - \gamma$  for simplicity. Let  $T_1$ ,  $T_2$  and  $T_3$  be the stopping times

$$\begin{aligned} T_1 &:= \inf\{n \geq m \text{ s.t. } Z_n(x - 3) > 1\}, \\ T_2 &:= \inf\{n \geq m \text{ s.t. } \alpha_n^-(x) \geq \gamma + \epsilon\} \\ T_3 &:= \inf\{n \geq m \text{ s.t. } Z_n(x - 2) \geq Ce^\epsilon Z_n(x)^{\gamma + \epsilon}\}. \end{aligned}$$

First, if  $Z_m(x - 2) \geq \text{Cst}(C, \epsilon)$ ,

$$\begin{aligned} \mathbb{P}(T_1 < T_2 \wedge T_3 | \mathcal{F}_m) &\leq \sum_{k=m}^{\infty} \mathbb{I}_{\{X_k = x-1\}} \frac{Ce^\epsilon Z_k(x)^{\gamma + \epsilon}}{Z_k(x - 2) + Z_k(x)} \frac{1}{Z_k(x - 1)} \\ &\leq 2Ce^\epsilon \sum_{k=Z_m(x-1)}^{\infty} k^{\gamma + \epsilon - 2} \leq \frac{2Ce^\epsilon (1 - (\gamma + \epsilon))^{-1}}{(Z_m(x - 1) - 1)^{1 - (\gamma + \epsilon)}} \end{aligned} \quad (13)$$

Second,

$$\mathbb{P}(T_2 < T_1 \wedge T_3 | \mathcal{F}_m) \leq \mathbb{P}(\mathcal{E}_{m, \epsilon/2}(x) | \mathcal{F}_m), \quad (14)$$

where  $\mathcal{E}_{\cdot, \cdot}(\cdot)$  is defined in (10). Indeed, assume  $\mathcal{E}_{m, \epsilon/2}(x)$  holds. For all  $n \geq m$ , let  $p(n)$  be the last time  $k \leq n$  s.t.  $\alpha_k^-(x) \leq \gamma$ . Then

$$Y_{p(n), n}(x - 2) \leq \gamma^{-1} \sum_{k=p(n)+1}^n \frac{\mathbb{I}_{\{X_k = x-2\}}}{Z_k(x)} \leq \gamma^{-1} (Ce^\epsilon)^{(\gamma + \epsilon)^{-1}} \sum_{k=p(n)+1}^n \frac{\mathbb{I}_{\{X_k = x-2\}}}{Z_k(x - 2)^{(\gamma + \epsilon)^{-1}}} \leq \frac{\epsilon}{2}$$

again if  $Z_m(x - 2) \geq \text{Cst}(C, \gamma, \epsilon)$ .

Third,

$$\mathbb{P}(T_3 < T_1 \wedge T_2 | \mathcal{F}_m) = \mathbb{E}[\mathbb{P}(T_3 < T_1 \wedge T_2 | \mathcal{F}_S) \mathbb{I}_{\{S < \infty\}}], \quad (15)$$



where

$$S := \inf\{n \geq m \text{ s.t. } Z_n(x-2) \geq CZ_n(x)^\gamma\}.$$

By Proposition 2.4 (d),

$$\mathbb{P}(T_3 < T_1 \wedge T_2 | \mathcal{F}_S) \leq \mathbf{Cst}(\epsilon)[Z_S(x-2)^{-1} + Z_S(x)^{-1}] \leq \mathbf{Cst}(C, \epsilon)Z_m(x)^{-\gamma}. \quad (16)$$

Now inequalities (13), (14), (15) and (16) enable us to conclude.  $\square$

### 2.3.3 Proof of Proposition 2.1

(a) For all  $n_0 \geq \mathbf{Cst}$ ,  $\epsilon > 0$ ,  $\alpha \in (0, 1)$ ,

$$\begin{aligned} \mathbb{P}(\{Z_{n_0+x}(x-2) = Z_{n_0+x}(x+2) = 2\} \cap \{Z_{n_0+x}(x-1) \in ((\alpha - \epsilon)n_0/2, (\alpha + \epsilon)n_0/2\} \\ \cap \{Z_{n_0+x}(x) \in (1/2 - \epsilon, 1/2 + \epsilon)n_0\}) \geq \mathbf{Cst}(n_0). \end{aligned}$$

Now apply Corollary 2.2 twice, with  $\pm := +$  and  $\pm := -$ , and Proposition 2.2.

(b) For all  $x \leq X_0$ , let us prove that  $\mathbb{P}(x-2 \notin R | x \in R) \geq \epsilon > 0$ , which will imply the conclusion. For all  $n_0 \in \mathbb{N}$ , let  $u_{n_0}(x)$  be the first time  $t$  such that  $Z_t(x) = n_0$ , and let  $t(x)$  be the time of first visit to  $x$ . Then

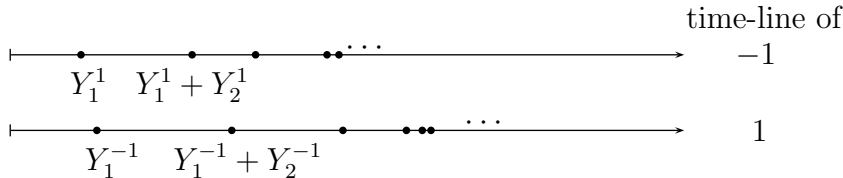
$$\mathbb{P}(\{Z_{u_{n_0}(x)}(x-2) = Z_{u_{n_0}(x)}(x-1) = 1\} \cap \{Z_{u_{n_0}(x)}(x+1) \geq n_0\}) \geq \mathbf{Cst}(n_0).$$

Now apply Corollary 2.2 with  $x := x$  and  $\pm := -$ .

## 3 Rubin continuous time-lines construction

### 3.1 W-urn

Again consider  $W$ -urn process with two colours  $-1$  and  $1$  studied in Section 1.1.1, defined as follows: start with  $Z_0(1)$  and  $Z_0(-1)$  balls of colours  $1$  and  $-1$  respectively. At each time step  $n \geq 0$ , pick a ball of colour  $i \in \{-1, 1\}$  in the urn with a probability  $W(Z_n(i))/(W(Z_n(-1)) + W(Z_n(1)))$ , and put it back together with a ball of the same colour.



We now construct a continuous-time process  $(\tilde{Z}_t(1), \tilde{Z}_t(-1))_{t \in \mathbb{R}_+}$  taking values in  $\mathbb{N}^2$ , which will be equal in law to  $(Z_n(1), Z_n(-1))_{n \in \mathbb{N}}$ , seen from the times of jumps. To this end we add balls of colour  $i \in \{-1, 1\}$  at rate  $W(Z_n(i))$ , using the following time-lines construction:

- Let  $(Y_k^{-1})_{k \in \mathbb{N}}$  and  $(Y_k^1)_{k \in \mathbb{N}}$  be two sequences of independent random variables of exponential law with  $\mathbb{E}Y_k^i = W(Z_0(i) + k - 1)^{-1}$ , independent from each other.
- Each of the colours 1 and  $-1$  has a clock with an alarm, set initially to  $Y_1^1$  and  $Y_1^{-1}$  respectively.
- Each time an alarm rings, we add a ball of the corresponding colour, say  $i$ . The other clock  $-i$  keeps running, while the new alarm with  $i$  is set at a time distance  $Y_{k+1}^i$  if  $k$  balls of colour  $i$  have already been added in the urn.

More precisely let, for all  $i \in \{-1, 1\}$  and  $t \in \mathbb{R}_+$ , let

$$S_1 := \left\{ \sum_{k=1}^n Y_k^i, n \in \mathbb{N} \right\}, S := S_1 \cup S_{-1},$$

let  $\xi_n$  be the  $n$ -th smallest element in  $S$ , with the convention that  $\xi_0 := 0$ . and let

$$\tilde{Z}_t(i) := \sup \left\{ l \geq 0 \text{ s.t. } \sum_{k=1}^l Y_k^i \leq t \right\} + Z_0(i).$$

**Lemma 3.1** *The processes  $(\tilde{Z}_{\xi_n}(1), \tilde{Z}_{\xi_n}(-1))_{n \geq 0}$  and  $(Z_n(1), Z_n(-1))_{n \geq 0}$  (from a  $W$ -urn) are equal in law.*

PROOF: By memoryless property of exponentials, for all  $n \in \mathbb{N}$ , the joint law of the first alarms after time  $\xi_n$  in the time-lines  $-1$  and  $1$  (conditioned on their past up to that time) is a couple of independent random variables of parameters  $W(\tilde{Z}_{\xi_n}(1))$  and  $W(\tilde{Z}_{\xi_n}(-1))$  respectively.

Now, if  $U$  and  $V$  are two independent random variables of parameters  $u$  and  $v$ , then  $\mathbb{P}[U < V] = u/(u + v)$ , which completes the proof.  $\square$

This result enables us to conclude that, if  $W$  is reciprocally summable, then only one of the balls is taken in the urn infinitely often.

**Proposition 3.1** *If  $\sum_{k \in \mathbb{N}} 1/W(k) < \infty$ , then  $Z_\infty(1) < \infty$  or  $Z_\infty(-1) < \infty$  a.s.*

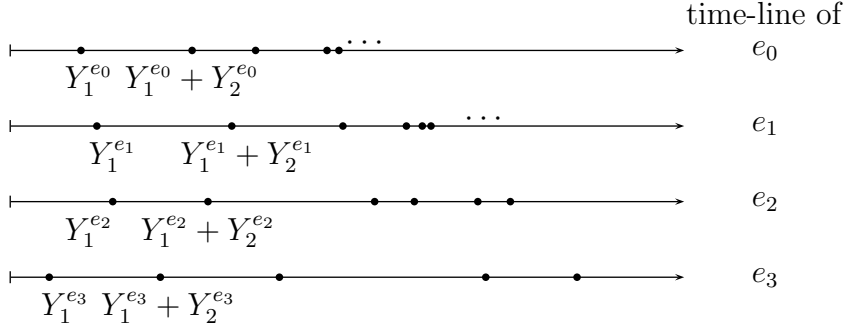
PROOF: We have

$$\begin{aligned} \mathbb{P}(Z_\infty(1) = Z_\infty(-1) = \infty) &= \mathbb{P}\left(\sum_{k=1}^{\infty} Y_k^1 = \sum_{k=1}^{\infty} Y_k^{-1}\right) \\ &= \mathbb{P}\left(Y_1^1 = \sum_{k=1}^{\infty} Y_k^{-1} - \sum_{k=2}^{\infty} Y_k^1\right) = 0, \end{aligned}$$

since  $Y_1^1$  has continuous density, independent from  $\sum_{k=1}^{\infty} Y_k^{-1} - \sum_{k=2}^{\infty} Y_k^1$ .  $\square$

### 3.2 ESIRW on a locally finite graph $(G, \sim)$

Let us similarly construct a continuous-time process  $(\tilde{X}_t)_{t \in \mathbb{R}_+}$ , which will be equal in law to  $(X_n)_{n \geq 0}$  for ESIRW, seen from times of jumps. Let  $E(G)$  be the set of (non-oriented) edges of  $(G, \sim)$ . The process starts at  $X_0 := x_0$  at time 0:



- Let  $(Y_k^e)_{e \in E(G), k \in \mathbb{N}}$  be a collection of independent exponential random variables with  $\mathbb{E}Y_k^e = W(k-1)^{-1}$ .
- Each edge  $e$  has its own clock, which only runs when the process  $(\tilde{X}_t)_{t \geq 0}$  is adjacent to  $e$ .
- Each time an edge  $e$  has just been crossed, and at time 0, its clock sets up an alarm at distance  $Y_{k+1}^e$  if  $e$  has been crossed  $k$  times so far ( $Y_1^e$  at time 0).
- Each time an edge  $e$  sounds an alarm,  $\tilde{X}_t$  crosses it instantaneously.

Let  $\tau_n$  be the  $n$ -th jump time of  $(\tilde{X}_t)_{t \geq 0}$ , with the convention that  $\tau_0 := 0$ .

**Lemma 3.2** (*Davis [4], Sellke [12]*) *The processes  $(\tilde{X}_{\tau_n})_{n \geq 0}$  and  $(X_n)_{n \geq 0}$  have the same distribution.*

The proof of Lemma 3.2 is left to the reader, being similar to that of 3.1.

Let

$$\mathcal{G}_\infty := \{e \in E(G) \text{ s.t. } Z_\infty(e) = \infty\}.$$

**Proposition 3.2** *If  $\sum_{n \in \mathbb{N}} 1/W(n) < \infty$ , then  $\mathcal{G}_\infty$  contains no even cycle.*

PROOF: For simplicity, let us denote an even cycle by  $\mathbb{Z}/l\mathbb{Z}$ ,  $l$  even. Let, for all  $i \in \mathbb{Z}/l\mathbb{Z}$ ,

$$T^i := \sum_{k=1}^{\infty} Y_k^{\{i, i+1\}}.$$

Then

$$\{\mathbb{Z}/l\mathbb{Z} \subseteq \mathcal{G}_\infty\} \subseteq \left\{ \sum_{x \in \mathbb{Z}/l\mathbb{Z}} (-1)^x T^x = 0 \right\}.$$

Now  $\sum_{x \in \mathbb{Z}/l\mathbb{Z}} (-1)^x T^x \neq 0$  a.s., which implies that

$$\mathbb{P}(\mathbb{Z}/l\mathbb{Z} \subseteq \mathcal{G}_\infty) = 0.$$

□

The technique carries over to show [12, 8] that, on graphs on bounded degree and if  $W$  is reciprocally summable, then  $\mathcal{G}_\infty$  is either a single edge or an odd cycle.

## 4 Short proof of a.s. localization of the VRRW on $\mathbb{Z}$ on five consecutive sites

Let, for all  $x \in \mathbb{Z}$ ,

$$\Omega(x) = \{\inf R' = x\}.$$

Then

$$\mathbb{P}(\cup_{x \in \mathbb{Z}} \Omega(x)) = 1,$$

by Proposition 2.1 (b), which asserts that the walk a.s. localizes on finitely many vertices. The aim of this section is to prove the following two propositions.

**Proposition 4.1** *For all  $x \in \mathbb{Z}$ ,  $\Upsilon(x) \subseteq \{Z_\infty(x-1) < \infty\} \cup \{Z_\infty(x+1) < \infty\}$  a.s.*

**Proposition 4.2** *For all  $x \in \mathbb{Z}$ ,  $\Omega(x) \subseteq \Upsilon(x+4)$  a.s.*

These will imply Theorem 2, i.e. a.s. localization on the VRRW on five consecutive vertices: a.s. on  $\Omega(x)$ ,  $Z_\infty(x+3) < \infty$  or  $Z_\infty(x+5) < \infty$  by Propositions 4.1 and 4.2, and the former would imply that  $R' = \{x, x+1, x+2\}$ , which holds with probability 0 by Proposition 2.2 (b).

We first propose an alternative time-lines construction for VRRWs in Section 4.1, which will enable us to couple two random walks in Section 4.2 with the following property: we will say that  $\mathcal{M}'$  is greater than  $\mathcal{M}$  if, at the time of  $n$ -th visit to any site  $x \in \mathbb{Z}$ ,  $\mathcal{M}'$  has more visited the right-hand side neighbour  $x+1$  than  $\mathcal{M}$ , whereas on the contrary  $\mathcal{M}$  has more visited  $x-1$  than  $\mathcal{M}'$ . Then we will prove Propositions 4.1 and 4.2 in Sections 4.3 and 4.4.

### 4.1 An alternative time-lines construction for VSIRWs on $\mathbb{Z}$

We introduce the following time-lines construction on directed edges  $\vec{E}(\mathbb{Z})$  of  $\mathbb{Z}$ , which will enable us to introduce naturally a coupling in Section 4.2. If  $e = (x, y) \in \vec{E}(\mathbb{Z})$ , let  $\underline{e} := x$ ,  $\bar{e} := y$ ,  $\sigma(e) := (y, x)$ .

The continuous time process  $(\tilde{X}_t)_{t \in \mathbb{R}_+}$  taking values in  $\mathbb{Z}$  will be defined as follows:

- Let  $(Y_k^e)_{e \in \vec{E}(\mathbb{Z}), k \in \mathbb{N}}$  be a collection of independent exponential random variables with expectation one.

- Each oriented edge  $e \in \vec{E}(\mathbb{Z})$  has its own clock, which only runs when the process  $(\tilde{X}_t)_{t \geq 0}$  is adjacent to  $e$ .
- Each time an edge  $e$  has just been crossed, the clock of  $\sigma(e)$  sets up an alarm at distance  $Y_{k+1}^{\sigma(e)}/W(\tilde{Z}_t(\underline{e}))$ , if  $\sigma(e)$  has been crossed  $k$  times so far. At time 0, we set up an initial alarm, at time distance  $Y_1^e$ , for the edges  $(x, x+1)$ ,  $x \geq x_0$ ,  $(x, x-1)$ ,  $x \leq x_0$ .
- Each time an edge  $e$  sounds an alarm,  $\tilde{X}_t$  crosses it instantaneously.

Let  $\tau_n$  be the  $n$ -th jump time of  $(\tilde{X}_t)_{t \geq 0}$ , with the convention that  $\tau_0 := 0$ .

**Lemma 4.1** *The processes  $(\tilde{X}_{\tau_n})_{n \geq 0}$  and  $(X_n)_{n \geq 0}$  have the same distribution.*

The proof of Lemma 4.1 is again left to the reader.

## 4.2 Coupling

Let us denote by  $\mathcal{M}$  the function which maps a (deterministic) “collection of alarms”  $\mathcal{Y} = (Y_k^e)_{e \in \vec{E}(\mathbb{Z}), k \in \mathbb{N}}$  and an initial site  $x_0$  to a continuous-time (deterministic) walk  $\mathcal{M}(\mathcal{Y}, x_0)$  on the vertices of  $\mathbb{Z}$ , as prescribed in Section 4.1.

**Definition 4.1** *Given  $\mathcal{Y} = (Y_k^e)_{e \in \vec{E}(\mathbb{Z}), k \in \mathbb{N}}$  and  $\mathcal{Y}' = ((Y'_k)^e)_{e \in \vec{E}(\mathbb{Z}), k \in \mathbb{N}}$  two collections of random variables on  $\mathbb{R}_+$ , we say that  $\mathcal{Y}' \gg \mathcal{Y}$  if, for all  $k \in \mathbb{N}$ ,  $x \in \mathbb{Z}$ ,  $(Y'_k)^{(x, x+1)} \leq Y_k^{(x, x+1)}$  and  $(Y'_k)^{(x, x-1)} \geq Y_k^{(x, x-1)}$  a.s.*

Given  $\mathcal{Y} = (Y_k^e)_{e \in \vec{E}(\mathbb{Z}), k \in \mathbb{N}}$  and  $\mathcal{Y}' = ((Y'_k)^e)_{e \in \vec{E}(\mathbb{Z}), k \in \mathbb{N}}$  two collections of random variables on  $\mathbb{R}_+$  we let, by a slight abuse of notation,  $\mathcal{M} = (\tilde{X}_t)_{t \in \mathbb{R}_+} := \mathcal{M}(\mathcal{Y}, x_0)$  and  $\mathcal{M}' = (\tilde{X}'_t)_{t \in \mathbb{R}_+} := \mathcal{M}(\mathcal{Y}', x_0)$  be the continuous-time random walks starting associated to  $\mathcal{Y}$  and  $\mathcal{Y}'$ .

For all  $i \in \mathbb{N}$ ,  $u > 0$ ,  $j \in \mathbb{Z}$  and  $e \in E(\mathbb{Z})$  (resp.  $\vec{e} \in \vec{E}(\mathbb{Z})$ ) non-oriented (resp. oriented) edge, let  $n_e(i)$  be the  $i$ -th visit time to  $e$ , let  $l_j(t)$  be the local time at  $j$  at time  $t$ , let  $t_j(u) := \inf\{t \geq 0 \text{ s.t. } l_j(t) = u\}$ , and let  $T_j$  be the total time spent in  $j$  for the random walk  $\mathcal{M}$ ; let  $n'_e(i)$  and  $T'_j$  be the similar notation for  $\mathcal{M}'$ .

For any non-oriented edge  $e = \{j, j+1\}$ , let  $\bar{e} := j+1$  and  $\underline{e} := j$ .

**Definition 4.2** *For all  $i \in \mathbb{N}$  and  $e \in E(\mathbb{Z})$ , let us define the property  $E_{i,e}$  as follows:*

$$Z'_{n'_e(i)}(\bar{e}) \geq Z_{n_e(i)}(\bar{e}) \text{ and } Z'_{n'_e(i)}(\underline{e}) \leq Z_{n_e(i)}(\underline{e}),$$

*with the convention that  $E_{i,j}$  holds whenever  $n_e(i) = \infty$  or  $n'_e(i) = \infty$ .*

**Lemma 4.2** *Assume  $\mathcal{Y}' \gg \mathcal{Y}$ , and  $W$  is nondecreasing. Then, for all  $i \in \mathbb{N}$  and  $e \in E(\mathbb{Z})$ ,  $E_{i,e}$  holds a.s.*

PROOF: Let, for all  $T > 0$ ,

$$\mathcal{P}_T := \{e \in E(\mathbb{Z}), i \in \mathbb{N} \text{ s.t. } n_e(i) \leq T \text{ and } n'_e(i) \leq T, E_{i,e} \text{ holds}\}.$$

Note that the property  $\mathcal{P}_T$  can only change on a discrete set of times; we prove it by induction. Assume that  $\mathcal{P}_T^-$  holds, i.e. that  $\mathcal{P}_t$  holds for all  $t < T$ . We want to deduce  $\mathcal{P}_T$ : assume for instance that  $n_e(i) = T$ ,  $n'_e(i) \leq T$ , with  $e = \{j, j+1\}$ ,  $j \geq x_0$ , and  $i$  odd, so that  $\tilde{X}_{n_e(i)} = \tilde{X}'_{n'_e(i)} = j+1$  (the other cases are similar). Obviously,

$$Z'_{n'_e(i)}(j+1) = Z'_{n'_e(i-1)}(j+1) + 1 \geq Z_{n_e(i)}(j+1) = Z_{n_e(i-1)}(j+1) + 1.$$

It remains to prove that

$$Z'_{n'_e(i)}(j) \leq Z_{n_e(i)}(j). \quad (17)$$

Let  $l_j := l_j(n_i(e)) - l_j(n_{i-1}(e))$  (resp.  $l'_j$ ) be the local time spent at  $j$  between times  $n_e(i-1)$  and  $n_e(i)$  (resp.  $n'_e(i-1)$  and  $n'_e(i)$ ). Then

$$l_j = \frac{Y_{(i+1)/2}^{(j,j+1)}}{W(Z_{n_e(i-1)}(j+1))} \geq \frac{(Y')_{(i+1)/2}^{(j,j+1)}}{W(Z'_{n'_e(i-1)}(j+1))} = l'_j.$$

Let

$$u := \inf\{0 \leq s \leq l'_j \text{ s.t. } Z'_{t'_j(n'_e(i)+s)}(j) > Z_{t_j(n_e(i)+s)}(j)\}.$$

Let us now consider the last jump occuring strictly before the local time at site  $j$  is  $u$ : it has to be a move from  $\mathcal{M}'$ , from  $j$  and  $j-1$  and back (possibly with a simultaneous move from  $\mathcal{M}$ ). At that last time, the numbers of visits to  $(j, j-1)$  are equal for  $\mathcal{M}$  and  $\mathcal{M}'$  (since those of  $j$  and  $(j, j+1)$  are) and, by  $\mathcal{P}_T^-$ , the numbers of visits to  $j-1$  is greater for  $\mathcal{M}$  than for  $\mathcal{M}'$ . Consequently  $\mathcal{M}$  must move before  $\mathcal{M}'$ , in  $j$ -th local time, after that last jump before  $u$ . Therefore  $u > l'_j$ , and  $\mathcal{P}_T$  holds.  $\square$

### 4.3 Proof of Proposition 4.1

Fix  $x \in \mathbb{Z}$ . Let  $\mathcal{Y} := (Y_k^e)_{e \in \vec{E}(\mathbb{Z}), k \in \mathbb{N}}$  be a collection of independent exponential random variables with expectation 1, and let

$$\mathcal{Y}'^{(n)} := ((Y'^{(n)})_k^e)_{e \in \vec{E}(\mathbb{Z}), k \in \mathbb{N}} := (Y_k^e + \mathbb{1}_{\{e=(x,x-1)\}} \mathbb{1}_{\{k=n\}})_{e \in \vec{E}(\mathbb{Z}), k \in \mathbb{N}}.$$

Let  $\mathcal{M} = (\tilde{X}_t)_{t \in \mathbb{R}_+} := \mathcal{M}(\mathcal{Y}, x_0)$  and  $\mathcal{M}'^{(n)} = (\tilde{X}'_t)_{t \in \mathbb{R}_+} := \mathcal{M}(\mathcal{Y}', x_0)$ , and let us use the notation from Section 4.2.

Let

$$\begin{aligned} \mathcal{Q} &:= \{Z_\infty(x+1) = Z_\infty(x-1) = \infty\} \cap \{T_x < \infty\}, \\ \mathcal{Q}'^{(n)} &:= \{Z'_\infty(x+1) = Z'_\infty(x-1) = \infty\} \cap \{T'_x < \infty\}. \end{aligned}$$

**Lemma 4.3** For all  $n \in \mathbb{N}$ ,  $\mathbb{P}(\mathcal{Q} \cap \mathcal{Q}'^{(n)}) = 0$ .

PROOF: If  $Z_\infty(x+1) = Z_\infty(x-1) = Z'_\infty(x+1) = Z'_\infty(x-1) = \infty$  and  $T_x < \infty$ ,  $T'_x < \infty$ , then

$$T_x = \sum_{k=1}^n \frac{Y_k^{(x,x+1)}}{Z_{n_{(x+1,x)}(k)}(x+1)} = \sum_{k=1}^n \frac{Y_k^{(x,x-1)}}{Z_{n_{(x-1,x)}(k)}(x-1)}$$

and, using Lemma 4.2,

$$T'_x = \sum_{k=1}^n \frac{(Y')_k^{(x,x+1)}}{Z'_{n'_{(x+1,x)}(k)}(x+1)} \leq T_x < \sum_{k=1}^n \frac{(Y')_k^{(x,x-1)}}{Z'_{n'_{(x-1,x)}(k)}(x-1)} = T'_x \text{ a.s.,}$$

which is contradictory. □

Let  $\mathcal{F}_n := \sigma(\tilde{X}_0, \dots, \tilde{X}_n) \subseteq \sigma(Y_k^e, 1 \leq k \leq n)$ . Then Lemma 4.3 implies

$$\mathbb{P}(\mathcal{Q}^c \cup (\mathcal{Q}'^{(n+1)})^c | \mathcal{F}_n) = 1.$$

But

$$\mathbb{P}(\mathcal{Q}^c | \mathcal{F}_n) \geq e^{-1} \mathbb{P}((\mathcal{Q}'^{(n+1)})^c | \mathcal{F}_n),$$

so that

$$\mathbb{P}(\mathcal{Q}^c | \mathcal{F}_n) \geq (1+e)^{-1}.$$

Now  $\mathbb{P}(\mathcal{Q}^c | \mathcal{F}_n) \xrightarrow[n \rightarrow \infty]{} \mathbb{I}_{\mathcal{Q}^c}$  a.s., so that  $\mathcal{Q}^c$  holds almost surely.

## 4.4 Proof of Proposition 4.2

By Corollary 2.1 (a)-(b), for all  $x \in \mathbb{Z}$ ,

$$\Omega(x) \subseteq \Upsilon(x) \subseteq \Upsilon(x) \cap \{\alpha_n^-(x+2) \xrightarrow[n \rightarrow \infty]{} \alpha_\infty^-(x+2) \in [0, 1)\} \subseteq \Upsilon(x+4) \cup \{\alpha_\infty^-(x+2) = 0\}.$$

Therefore, if we let

$$\bar{\Omega}(x) := \Omega(x) \cap \{\alpha_\infty^-(x+2) = 0\},$$

it is sufficient to show that  $\mathbb{P}(\bar{\Omega}(x)) = 0$ .

Let us assume  $x := 0$  for simplicity. Our proof relies on the study of the asymptotic behaviour of  $\log Z_n(3)/Z_n(2)$ : roughly speaking, we will show that this quantity must converge to 0 on  $\bar{\Omega}(0)$  but that, on the other hand, its convergence to 0 can only happen with zero probability, due to unstability.

For all  $n \in \mathbb{N}$  and  $x \in \mathbb{Z}$ , let

$$R_n := Z_n(4) + Z_n(2) - Z_n(1) - Z_n(3),$$

$$z_n := \log \frac{Z_n(3)}{Z_n(2)}, \quad y_n := \frac{R_n}{Z_n(2)Z_n(3)},$$

and let  $t_n(x)$  (resp.  $t_n^\pm(x)$ ) be the  $n$ -th visit time of  $x$  (resp.  $(x, x \pm 1)$ , counted once the edge has been visited), and let  $Z_n^\pm(x)$  be the number of visits of the edge  $(x, x \pm 1)$  at time  $n$ .

Note that, for all  $i \in \mathbb{N}$ ,

$$R_i = Z_i^-(5) - Z_i^+(0) + \mathbb{1}_{\{X_i=2 \text{ or } X_i \geq 4\}} + \text{Cst}(x_0), \quad (18)$$

so that

$$R_{t_i^+(2)} = Z_{t_i^+(2)}^+(4) - Z_{t_i^+(2)}^-(1) + \text{Cst}(x_0). \quad (19)$$

For all  $\alpha \in (0, 1)$  and  $k \geq j$ , we write  $j \leftrightarrow_\alpha k$  if, for all  $i \in [j, k]$ ,  $\alpha_i^-(2) \vee \alpha_i^+(3) \leq \alpha$ .

**Lemma 4.4** *Assume  $Z_n(2) \geq \text{Cst}$ . Then there exist  $\alpha_0 := \text{Cst}$ ,  $(\eta_{j,k})_{k \geq j \geq n}$  and  $(r_{j,k})_{k \geq j \geq n}$  such that, and  $k \geq j \geq n$  with  $j \leftrightarrow_{\alpha_0} k$ ,*

$$z_k - z_j = \sum_{i=j+1}^k \mathbb{1}_{\{X_{i-1}=2, X_i=3\}} y_i + \eta_{n,k} + r_{n,k}, \quad (20)$$

where, for all  $\alpha \leq \alpha_0$  and  $\epsilon > 0$ ,

$$\mathbb{P} \left( \sup_{k \geq j \geq n} |\eta_{j,k}| \mathbb{1}_{j \leftrightarrow_\alpha k} \geq \epsilon \mid \mathcal{F}_n \right) \leq \text{Cst} \frac{\alpha}{\epsilon^2 Z_n(2)}, \quad |r_{j,k}| \leq \frac{\text{Cst}}{Z_j(2)}. \quad (21)$$

Lemma 4.4 is proved in Section 4.4.1.

**Lemma 4.5**  $\bar{\Omega}(0) \subseteq \{\limsup(\alpha_n^+(3)/\alpha_n^-(2)) \leq 1\} \cap \{\sum |y_{t_n^+(2)}| \mathbb{1}_{t_n^+(2) < \infty} < \infty\}$ .

Lemma 4.5 is proved in Section 4.4.2.

For simplicity we will use the notation  $t_n := t_n^+(2)$  until the beginning of Section 4.4.1. Let  $E_1 := \{(2, 1), (0, 1)\}$  be the set of edges pointing to 1. We introduce the continuous-time construction of Section 4.1 in order to analyze the walk after time  $t_n$ , but modify the rule for the two edges in  $E_1$ : each time an edge  $e \in \sigma(E_1)$  has been crossed, we set up an alarm on  $\sigma(e)$  at time distance  $V_{k+1}$  if site 1 has been visited  $k$  times after time  $n$ . The walk thus defined depends on the sequences  $(Y_k^e)_{e \in \bar{E}(\mathbb{Z}) \setminus E_1}$ ,  $(V_k)_{k \in \mathbb{N}}$  and on the starting point  $X_{t_n}$ .

Let  $(Y_k^e)_{e \in \bar{E}(\mathbb{Z}) \setminus E_1}$  (resp.  $((Y'_k)^e)_{e \in \bar{E}(\mathbb{Z}) \setminus E_1}$ ) be collections of independent exponential random variables with expectation one, and let  $(V_k)_{k \in \mathbb{N}}$  (resp.  $(V'_k)_{k \in \mathbb{N}}$ ) be sequences of independent exponential random variables with expectation  $\mathbb{E} V_k = (n_0 + k - 1)^{-1}$  (resp.  $\mathbb{E} V'_k = (n_0 + k - 1)^{-1}$ ); these define two continuous-time walks  $\mathcal{M}$  and  $\mathcal{M}'$ . We choose  $n_0 := Z_{t_n}(1)$  and  $n'_0 := Z_{t_n}(1) - a\sqrt{Z_{t_n}(1)}$ .

Note that both  $\mathcal{M}$  and  $\mathcal{M}'$  are continuous-time versions of a VRRW:  $\mathcal{M}$  of the original walk  $(X_k)_{k \geq t_n}$ , and  $\mathcal{M}'$  of the VRRW  $:= (X'_k)_{k \geq t_n}$  taken after discounting the number of visits to 1 of  $a\sqrt{Z_{t_n}(1)}$ .



Given  $\epsilon > 0$  and  $n \in \mathbb{N}$ , assume  $\alpha_{t_n}^-(2) < \epsilon$ , and let

$$T_n := \inf\{k \geq t_n \text{ s.t. } \alpha_k^+(3) \vee \alpha_k^-(2) \geq (1 + \epsilon)\alpha_n^-(2) \\ \text{or } \sum_{i=k+1}^n |y_{i-1}| \mathbb{1}_{\{X_{i-1}=2, X_i=3\}} > \epsilon \text{ or } X_k = 0\}.$$

We let  $Z'(\cdot)$ ,  $z'$ ,  $R'$ ,  $T'$ ,  $\overline{\Omega}'(0)$  and  $t'$  be the notation for  $\mathcal{M}'$  equivalent to the one already defined for  $\mathcal{M}$ .

Let

$$\mathcal{Q}_n := \overline{\Omega}(0) \cap \{T_n = \infty\}, \\ \mathcal{Q}'_n := \overline{\Omega}'(0) \cap \{T'_n = \infty\}.$$

**Lemma 4.6** *If  $\epsilon \leq \text{Cst}$ ,  $Z_{t_n}(1) \geq \text{Cst}$  and  $a \geq \text{Cst}$ , then  $\mathbb{P}((\mathcal{Q}_n \cap \mathcal{Q}'_n)^c | \mathcal{F}_{t_n}) \geq \text{Cst}$ .*

Lemma 4.6 is proved in Section 4.4.3.

**Lemma 4.7** *For all  $\delta > 0$ , there exists  $C(a, \delta)$  (depending only on  $a$  and  $\delta$ ) such that*

$$\mathbb{P}((\mathcal{Q}'_n)^c | \mathcal{F}_{t_n}) \leq C(a, \delta) \mathbb{P}(\mathcal{Q}_n^c | \mathcal{F}_{t_n}) + \delta.$$

Lemma 4.7 completes the proof of Proposition 4.4: indeed, using also Lemma 4.6, for all  $n \in \mathbb{N}$ ,

$$\text{Cst} \leq \mathbb{P}(\mathcal{Q}_n^c | \mathcal{F}_{t_n}) + \mathbb{P}((\mathcal{Q}'_n)^c | \mathcal{F}_{t_n}) \leq (1 + C(a, \delta)) \mathbb{P}(\mathcal{Q}_n^c | \mathcal{F}_{t_n}) + \delta,$$

which implies  $\mathbb{P}(\mathcal{Q}_n^c | \mathcal{F}_{t_n}) \geq \text{Cst}(a)$  if we choose  $\delta \leq \text{Cst}$ . Now  $\mathbb{P}(\mathcal{Q}_n^c | \mathcal{F}_{t_n}) \xrightarrow[n \rightarrow \infty]{} \mathbb{1}_{\mathcal{Q}_n^c}$  a.s., so that  $\mathcal{Q}_n^c$  holds a.s. Now  $\liminf_{n \rightarrow \infty} \mathcal{Q}_n^c = \overline{\Omega}(0)^c$ , since  $T_n = \infty$  for some  $n$  on  $\overline{\Omega}(0)$ , using Lemma 4.5 and Corollary 2.1 (a).

PROOF: There is a one-to-one correspondence between  $(V_k)_{k \in \mathbb{Z}}$  and a simple birth process  $\{N_t, t \geq 0\}$  with initial population size  $n_0$ , defined by

$$N_t := n_0 + \sup \left\{ k \in \mathbb{N} \text{ s.t. } \sum_{i=1}^k V_i \leq t \right\}.$$

By a result of D. Kendall [7],  $\{N_{\log(1+t/W)}, t \geq 0\}$  is a Poisson process with unit parameter, where  $W := \lim N_t e^{-t}$  is a Gamma random variable  $\Gamma(n_0, 1)$  with shape  $n_0$  and scale 1.

The same remark applies to  $(V'_k)_{k \in \mathbb{Z}}$ . Now recall that  $U \sim \Gamma(\lambda, 1)$  has density  $\phi_\lambda(x) := x^{\lambda-1} e^{-x} / \Gamma(\lambda)$  w.r.t. Lebesgue measure  $\mathcal{L}(dx)$ , expectation and variance  $\lambda$ .

Assume  $\lambda \geq 1$ ; for all  $a, c > 0$ , there exist  $c_1, c_2 := \mathbf{Cst}(a, c)$  such that, for all  $x$  such that  $|x - \lambda| \leq c\sqrt{\lambda}$ ,  $\phi_\lambda(x)/\phi_{\lambda-\sqrt{\lambda}a}(x) \in [c_1, c_2]$ . Hence, for any Borel subset  $A$  of  $\mathbb{R}$ ,

$$\begin{aligned} \mathbb{P}(W' \in A) &\leq \mathbb{P}(|W' - n'_0| \geq \sqrt{n'_0}c) + \mathbb{P}(W' \in A \cap [n'_0 - \sqrt{n'_0}c, n'_0 + \sqrt{n'_0}c]) \\ &\leq c^{-2} + \mathbf{Cst}(a, b)\mathbb{P}(W \in A). \end{aligned}$$

We conclude by choosing  $c := \delta^{-1/2}$ . □

#### 4.4.1 Proof of Lemma 4.4

Assume  $j \leftrightarrow_{\alpha_0} k$ , with  $\alpha_0 \leq \mathbf{Cst}$ . Let us first estimate  $h(Z_k(3)) - h(Z_j(3))$ : using Proposition 2.4 (a), there exists a constant  $C > 0$  such that

$$h(Z_k(3)) = Y_k^+(2) + Y_k^-(4) = Y_k^+(2) + U_{k,+}^+ - \frac{\mathbb{I}_{\{X_k \geq 4\}}}{Z_{k-1}(3)} + C$$

But

$$U_{k,+}^+(3) = U_{k,+}^-(3) + \check{U}_{k,+}(3).$$

Now we estimate  $U_{k,+}^-(3)$ . We assume  $X_0 \leq 2$  (the other case is similar): then, for all  $i \in \mathbb{N}$ ,  $t_i^+(2) < t_i^-(3) < t_{i+1}^+(2)$ , and  $Z_{t_i^-(3)}(x) = Z_{t_{i+1}^+(2)}(x)$  for  $x = 3, 4$ , so that

$$\begin{aligned} U_{k,+}^-(3) - U_{j,+}^-(3) &= \sum_{i=Z_n^-(3)+1}^{Z_k^-(3)} \frac{Z_{t_i^-(3)}(4)}{Z_{t_i^-(3)}(3)Z_{t_i^-(3)}(2)} \\ &= \sum_{i=Z_j^-(3)+1}^{Z_k^-(3)} \left( \frac{Z_{t_{i+1}^+(2)}(4)}{Z_{t_{i+1}^+(2)}(3)Z_{t_{i+1}^+(2)}(2)} + \frac{Z_{t_{i+1}^+(2)}(4)}{Z_{t_{i+1}^+(2)}(3)} \left( \frac{1}{Z_{t_i^-(3)}(2)} - \frac{1}{Z_{t_{i+1}^+(2)}(2)} \right) \right) \\ &= \sum_{l=j+1}^k \frac{\mathbb{I}_{\{X_{l-1}=2, X_l=3\}}}{Z_{l-1}(2)} \frac{Z_{l-1}(4)}{Z_{l-1}(3)} + r_{j,k}^1, \end{aligned}$$

where

$$|r_{j,k}^1| \leq 2 \sup_{j \leq l < k} \left( \frac{Z_l(4)}{Z_l(3)} \right) \frac{1}{Z_j(3)} \leq \frac{1}{2Z_j(3)}$$

if  $\alpha \leq \mathbf{Cst}$ .

In summary,

$$\log \frac{Z_k(3)}{Z_j(3)} = \sum_{l=j+1}^k \frac{\mathbb{I}_{\{X_{l-1}=2, X_l=3\}}}{Z_{l-1}(2)Z_{l-1}(3)} (Z_{l-1}(2) + Z_{l-1}(4)) + \check{U}_{k,+}(3) - \check{U}_{j,+}(3) + r_{j,k}^2,$$

where  $|r_{j,k}^2| \leq 3/Z_j(2)$ , using (12).

Similarly,

$$\log \frac{Z_k(2)}{Z_j(2)} = \sum_{l=j+1}^k \frac{\mathbb{I}_{\{X_{l-1}=2, X_l=3\}}}{Z_{l-1}(2)Z_{l-1}(3)} (Z_{l-1}(1) + Z_{l-1}(3)) + \check{U}_{k,-}(2) - \check{U}_{j,-}(2) + r_{j,k}^3,$$

where  $|r_{j,k}^3| \leq 3/Z_j(2)$ .

Now, if  $\alpha_0 \leq \text{Cst}$ , then

$$\sum_{l=j}^k |y_{l-1} - y_l| \mathbb{I}_{\{X_{l-1}=2, X_l=3\}} \leq \sum_{l=j}^k \frac{\mathbb{I}_{\{X_{l-1}=2, X_l=3\}}}{Z_l(2)Z_l(3)} \leq \frac{2}{Z_j(2)}.$$

This provides (20), letting

$$\eta_{j,k} := \check{U}_{k,+}(3) - \check{U}_{j,+}(3) - (\check{U}_{k,-}(2) - \check{U}_{j,-}(2)).$$

Let us now estimate  $\eta_{j,k}$ . The following Lemma 4.8 will enable us to conclude, using Chebyshev's inequality.

For all  $n \in \mathbb{N}$ ,  $x \in \mathbb{Z}$  and  $\alpha > 0$ , let

$$\begin{aligned} \check{U}_{n,\pm}(x) &:= \hat{U}_{n,\pm}^+(x) - \hat{U}_{n,\pm}^-(x) \\ \tilde{U}_{n,\pm}(x) &:= \sum_{k=1}^n (\check{U}_{k,\pm}(x) - \check{U}_{k-1,\pm}(x)) \mathbb{I}_{\{\alpha_{k-1}^\pm(x) \leq \alpha\}} \end{aligned}$$

**Lemma 4.8** *For all  $x \in \mathbb{Z}$ ,  $k \geq n$ ,  $\alpha \leq \text{Cst}$ ,*

$$\mathbb{E} \left( \sup_{k \geq n} (\tilde{U}_{k,\pm}(x) - \tilde{U}_{n,\pm}(x))^2 | \mathcal{F}_n \right) \leq \text{Cst} \alpha Z_n(x)^{-1}.$$

PROOF:  $(\tilde{U}_{n,\pm}(x))_{n \geq 0}$  is a martingale, and

$$\mathbb{E}((\tilde{U}_{k+1,\pm}(x) - \tilde{U}_{k,\pm}(x))^2 | \mathcal{F}_k) \leq \frac{\mathbb{I}_{\{X_k=x\}}}{Z_k(x)^2} \left( \alpha + \frac{\alpha^2}{1-\alpha} \right) \leq 2\alpha \frac{\mathbb{I}_{\{X_k=x\}}}{Z_k(x)^2},$$

which enables us to conclude by Doob's inequality. □

#### 4.4.2 Proof of Lemma 4.5

Given  $\epsilon > 0$  and  $n \in \mathbb{N}$ , let

$$T_n := \inf \{k \geq n \text{ s.t. } \alpha_k^-(2) \geq (1 + \epsilon^2) \alpha_n^-(2) \text{ or } Y_k^+(0) - Y_n^+(0) > \epsilon^2\}.$$

Let us first prove that, for all  $\epsilon > 0$  and sufficiently large  $n \in \mathbb{N}$ , if

$$\alpha_n^+(3) \geq (1 + \epsilon)^4 \alpha_n^-(2) \text{ with } \alpha_n^-(2) \leq \epsilon^2 \leq \text{Cst}, \quad (22)$$

then

$$\mathbb{P}(\{T_n < \infty\} \cup \{\liminf \alpha_k^-(3) > 0\} | \mathcal{F}_n) \geq 1/2. \quad (23)$$

This will imply the first part of the inclusion. Indeed let, for all  $p \in \mathbb{N}$ ,

$$\begin{aligned} \mathcal{A} &:= \{\limsup_{k \rightarrow \infty} (\alpha_k^+(3) / \alpha_k^-(2)) \leq 1\}, \\ \mathcal{B}_p &:= \{T_p < \infty\} \cup \{\liminf_{k \rightarrow \infty} \alpha_k^+(3) > 0\} \cup \{\limsup_{k \rightarrow \infty} \alpha_k^-(2) > 0\}. \end{aligned}$$

Then (23) implies, for all  $p \in \mathbb{N}$ , that for large  $n \in \mathbb{N}$ ,  $1/2 \leq \mathbb{P}(\mathcal{A} \cup \mathcal{B}_p | \mathcal{F}_n)$ , which converges a.s. to  $\mathbb{I}_{\mathcal{A} \cup \mathcal{B}_p}$  as  $n \rightarrow \infty$ , so that  $\mathcal{A} \cup \mathcal{B}_p$  holds a.s.; subsequently  $\mathcal{A} \cup \liminf \mathcal{B}_p$  holds a.s. Now, by Proposition 2.4 (c) and Corollary 2.1 (a) and (d),

$$\overline{\Omega}(0) \cap \{\liminf \alpha_k^+(3) > 0\} \subseteq \overline{\Omega}(0) \cap \Upsilon(-1, 0, 1) = \emptyset,$$

which implies  $\overline{\Omega}(0) \subseteq \limsup \mathcal{B}_p^c \subseteq \mathcal{A}$  and enables to conclude.

Let us assume (22), and  $\alpha_n^+(3) \leq \epsilon^2$  w.l.o.g (possibly by choosing  $n$  larger), and show (23). Fix  $\alpha := (1 + \epsilon)^4 \alpha_n^-(2)$ , and let  $p$  be the largest  $i \in [n + 1, k]$  such that  $\alpha_{i-1}^-(2) > \alpha$  or  $\alpha_{i-1}^+(3) > \alpha$ ; then  $\alpha_p^+(3) \geq (1 + \epsilon)^2 \alpha_p^-(2)$ . Note that, for all  $i \geq p$ ,  $Z_i(2)/Z_i(3) \geq \alpha_i^-(3) \geq 1 - \epsilon^2$ , since  $Z_i(3) \leq Z_i(2) + Z_i(4)$ .

Let  $t_j := t_j^+(2)$  for simplicity. In order to make use of Lemma 4.4, we need to estimate  $\sum_{j=Z_p^+(2)+1}^{Z_k^+(2)} y_{t_j}$ . Now, for all  $i \geq p$ , using (18),

$$R_i \geq R_p - (Z_i^+(0) - Z_p^+(0)) - 1. \quad (24)$$

On one hand,

$$\begin{aligned} \sum_{j=Z_p^+(2)+1}^{Z_k^+(2)} \frac{1}{Z_{t_j}(2)Z_{t_j}(3)} &\geq \sum_{j=0}^{Z_k^+(2)-Z_p^+(2)-1} \frac{1}{(Z_p(2)+j)(Z_p(2)+j+1)} (1 - \text{Cst} \cdot \alpha) \\ &\geq \left( \frac{1}{Z_p(2)} - \frac{1}{Z_k(2)} \right) (1 - \text{Cst} \cdot \alpha). \end{aligned} \quad (25)$$

On the other hand, assume  $Z_n(2) \geq \text{Cst}$ , so that  $Z_i(2)Z_i(3) \geq (1 - \epsilon^2)^2 Z_i(2)(Z_i(2) + 1)$ ; then

$$\begin{aligned} \Delta_{p,k} &:= \sum_{i=p}^{k-1} \frac{Z_i^+(0) - Z_p^+(0)}{Z_i(2)Z_i(3)} \mathbb{I}_{\{X_i=2\}} \leq (1 - \epsilon^2)^{-2} \sum_{i=p}^{k-1} (Z_i^+(0) - Z_p^+(0)) \left( \frac{1}{Z_{i-1}(2)} - \frac{1}{Z_i(2)} \right) \\ &\leq (1 - \epsilon^2)^{-2} \sum_{i=p}^{k-1} \frac{\mathbb{I}_{\{X_{i-1}=0, X_i=1\}}}{Z_{i-1}(2)} \leq (1 - \epsilon^2)^{-4} \alpha \epsilon^2 \end{aligned} \quad (26)$$

Now, if  $Z_n(1) \geq \text{Cst}(\epsilon)$ , using (21),

$$\mathbb{P}\left(\sup_{k \geq j \geq n} |\eta_{j,k}| \mathbb{I}_{j \leftrightarrow \alpha k} \mid \mathcal{F}_n\right) \geq 1 - \frac{\text{Cst}}{\epsilon^4 \alpha_n^-(2)} \left( \frac{1}{Z_n(2)} + \frac{1}{Z_n(3)} \right) \geq \frac{1}{2},$$

where we use  $\alpha_n^-(2)Z_n(3) = \alpha_n^+(2)Z_n(1) \geq (1 - \epsilon^2)Z_n(1)$ .

Now assume that  $E$  holds. If  $Z_n(1) \geq \text{Cst}(\epsilon)$ , then Lemma 4.4 implies, together with (25) and (26),

$$\log \frac{Z_k(3)}{Z_k(2)} - \log \frac{Z_p(3)}{Z_p(2)} + \text{Cst}\epsilon^2\alpha \geq R_p \left( \frac{1}{Z_p(3)} - \frac{1}{Z_k(3)} \right) \geq \frac{R_p}{Z_p(3)} - \frac{R_k + Z_k(1)}{Z_k(3)},$$

where we use in the second inequality that  $R_k \geq R_p - Z_k(1)$ , by (24).

Therefore

$$\frac{Z_k(4)}{Z_k(2)} \geq \log \frac{Z_k(3)}{Z_k(2)} + \frac{R_k + Z_k(1)}{Z_k(3)} - \text{Cst}\epsilon^2\alpha \geq \frac{Z_p(4) - Z_p(1)}{Z_p(2)} - \text{Cst}\epsilon^2\alpha \geq \frac{\epsilon\alpha}{2}$$

Therefore  $\liminf \alpha_k^+(3) > 0$  on  $E$ , which completes the proof of the first part of the inclusion.

Now assume  $\overline{\Omega}(0) \cap \{\limsup(\alpha_k^+(3)/\alpha_k^-(2)) \leq 1\}$  holds. Then  $Z_n(3)/Z_n(2)$  converges to 1, and the estimates (24), (25) and (26) ensure that  $\sum y_{i-1}^- \mathbb{I}_{\{X_{i-1}=2, X_i=3\}} < \infty$ . On the other hand Lemma 4.4 ensures, together with Lemma 4.8 and Doob's martingale Theorem that  $\limsup_{k,n \rightarrow \infty: k \geq n} |\sum y_{i-1} \mathbb{I}_{\{X_{i-1}=2, X_i=3\}}| < \infty$ , which completes the second part of the inclusion.

#### 4.4.3 Proof of Lemma 4.6

Let  $\alpha := (1 + \epsilon)\alpha_n^-(2)$ , and let  $k \geq j \geq n$  be such that  $t_k \leq T$ ,  $t'_k \leq T'$ .

Lemma 4.4 still holds for  $\mathcal{M}'$ , with the modification that  $Z_i(1)$  has to be replaced by  $Z_i(1) - a\sqrt{Z_{t_n}(1)}$  in  $y_i$ : letting  $\eta'_{\cdot,\cdot}$ ,  $\leftrightarrow'_\alpha$  and  $r'_{\cdot,\cdot}$  be the corresponding notation,

$$z'_{t'_k} - z_{t_k} = z'_{t'_j} - z_{t_j} + \sum_{i=j+1}^k \frac{a\sqrt{Z_{t_n}(1)}}{Z'_{t'_i}(2)Z'_{t'_i}(3)} + \sum_{i=j+1}^k (y'_{t'_i} - y_{t_i}) + \eta'_{t'_j, t'_k} - \eta_{t_j, t_k} + r'_{t'_j, t'_k} - r_{t_j, t_k}.$$

Let us first estimate  $y'_{t'_i} - y_{t_i}$ , for all  $i \in [n, k]$ : the coupling of Lemma 4.2 holds, as long as  $n_e(i) \leq T$ ,  $n'_e(i) \leq T'$ . Therefore  $Z'_{t'_i}(3) \geq Z_{t_i}(3)$  (and thus  $z'_{t'_k} \geq z_{t_k}$ ), thus  $Z'^+_{t'_i}(3) \geq Z^+_{t_i}(3)$  (since the numbers of visits of edge  $(2, 3)$  are equal for both walks at those times), which implies subsequently  $Z'^+_{t'_i}(4) \geq Z^+_{t_i}(4)$ . Similarly  $Z'^-_{t'_i}(1) \leq Z^-_{t_i}(1)$  so that, using identity 19,  $R'_{t'_i} \geq R_{t_i}$ .

On the other hand, using that  $|x| \leq 2|\ln(1+x)|$  for  $|x| < 1$ ,

$$\left| \frac{Z_{t_i}(2)Z_{t_i}(3)}{Z'_{t'_i}(2)Z'_{t'_i}(3)} - 1 \right| \leq 2 \left| \ln \frac{Z_{t_i}(2)Z_{t_i}(3)}{Z'_{t'_i}(2)Z'_{t'_i}(3)} \right| \leq 2 \left( \ln \frac{Z'_{t'_i}(3)}{Z_{t_i}(3)} + \ln \frac{Z_{t_i}(2)}{Z'_{t'_i}(2)} \right) = 2(z'_{t'_i} - z_{t_i}).$$

Hence

$$y'_{t'_i} - y_{t_i} \geq -|R_{t_i}| \left| \frac{1}{Z_{t_i}(2)Z_{t_i}(3)} - \frac{1}{Z'_{t'_i}(2)Z'_{t'_i}(3)} \right| \geq -2|y_{t_i}|(z'_{t'_i} - z_{t_i}).$$

Now let

$$u(k) := \operatorname{argmax}_{i \in [n, k]} (z'_{t'_i} - z_{t_i}), \quad \tau_k := z'_{t'_{u(k)}} - z_{t_{u(k)}}.$$

Using (25) in the proof of Lemma 4.5, and  $\sum_{i=j+1}^k |y_{t_i}| \leq \epsilon$ , we obtain that

$$\begin{aligned} z'_{t'_k} - z_{t_k} &\geq z'_{t'_j} - z_{t_j} - 2\epsilon\tau_k + a\sqrt{Z_{t_n}(1)}(Z_{t_j}(2)^{-1} - Z_{t_k}(2)^{-1})(1 - \operatorname{Cst}\alpha) \\ &\quad + \eta'_{t'_j, t'_k} - \eta_{t_j, t_k} + r'_{t'_j, t'_k} - r_{t_j, t_k}. \end{aligned} \quad (27)$$

Given  $b > 0$ , let

$$\Delta := \left\{ \sup_{k \geq j \geq n} |\eta_{t_j, t_k}| \mathbb{I}_{j \leftrightarrow_{\alpha} k} \leq b \frac{\sqrt{Z_{t_n}(1)}}{Z_{t_n}(2)} \right\} \cap \left\{ \sup_{k \geq j \geq n} |\eta'_{t'_j, t'_k}| \mathbb{I}_{j \leftrightarrow'_{\alpha} k} \leq b \frac{\sqrt{Z_{t_n}(1)}}{Z_{t_n}(2)} \right\}.$$

Then, by (21) and Chebyshev's inequality, if  $b \geq \operatorname{Cst}$ , then

$$\mathbb{P}(\Delta \mid \mathcal{F}_{t_n}) \geq 1/2.$$

Now assume  $\Delta$  holds, and apply (27) for  $j := n$  and  $k$  such that  $Z_{t_k}(2) \geq 2Z_{t_n}(2)$ : if  $Z_{t_n}(1) \geq \operatorname{Cst}(b)$ , then

$$\tau_k \geq \left( \frac{a}{2} - 3b \right) \frac{\sqrt{Z_{t_n}(1)}}{Z_{t_n}(2)} - 2\epsilon\tau_k. \quad (28)$$

Apply again (27) for  $j := u(k)$  and  $k$ , and use (28):

$$z'_{t'_k} - z_{t_k} \geq (1 - 2\epsilon)\tau_k - 3b \frac{\sqrt{Z_{t_n}(1)}}{Z_{t_n}(2)} \geq \left\{ \left( \frac{1 - 2\epsilon}{1 + 2\epsilon} \right) \left( \frac{a}{2} - 3b \right) - 3b \right\} \frac{\sqrt{Z_{t_n}(1)}}{Z_{t_n}(2)} \geq \frac{a}{3} \frac{\sqrt{Z_{t_n}(1)}}{Z_{t_n}(2)},$$

if  $a \geq \operatorname{Cst}(b)$  and  $\epsilon \leq \operatorname{Cst}$ .

We conclude by the remark that  $z_{t_n}$  (resp.  $z'_{t'_n}$ ) converge to 0 on  $\overline{\Omega}(0)$  (resp.  $\overline{\Omega}'(0)$ ), so that  $\mathbb{P}(\mathcal{Q}_n \cap \mathcal{Q}'_n \cap \Delta) = 0$ , and  $\mathbb{P}((\mathcal{Q}_n \cap \mathcal{Q}'_n)^c \mid \mathcal{F}_{t_n}) \geq \mathbb{P}(\Delta \mid \mathcal{F}_{t_n}) \geq 1/2$  if  $b \geq \operatorname{Cst}$ .

## 5 Appendix

We recall Lévy's conditional Borel-Cantelli Lemma. Let  $\mathbb{F} = (\mathcal{F}_n)_{n \in \mathbb{N}}$  be a filtration, and let  $(\xi_n)_{n \in \mathbb{N}}$  be an  $\mathbb{F}$ -adapted sequence taking values in  $\mathbb{R}_+$ .

**Lemma 5.1** *Assume  $(\xi_n)_{n \in \mathbb{N}}$  is a.s. bounded by a constant  $C > 0$ . Then*

$$\left\{ \sum_{k=1}^{\infty} \xi_k < \infty \right\} = \left\{ \sum_{k=1}^{\infty} \mathbb{E}(\xi_k \mid \mathcal{F}_{k-1}) < \infty \right\}.$$

PROOF: Let

$$M_n := \sum_{k=1}^n (\xi_k - \mathbb{E}(\xi_k | \mathcal{F}_{k-1})),$$

with  $\mathcal{F}_0 := \emptyset$ . Then  $(M_n)_{n \in \mathbb{N}}$  is a martingale, and

$$\langle M \rangle_n := \sum_{k=1}^{n-1} \mathbb{E}((M_k - M_{k-1})^2 | \mathcal{F}_{k-1}) \leq \sum_{k=1}^{n-1} \mathbb{E}(\xi_k^2 | \mathcal{F}_{k-1}) = O\left(\sum_{k=1}^{n-1} \mathbb{E}(\xi_k | \mathcal{F}_{k-1})\right).$$

Now, almost surely, either  $\langle M \rangle_\infty < \infty$  or  $M_n / \langle M \rangle_n \xrightarrow[n \rightarrow \infty]{} 0$ , by Lévy's strong law of large numbers (see for instance [15], Chapter 12, § 14), which yields the result.  $\square$

**Acknowledgment.** I would like to thank Terry Lyons for encouraging me to give a graduate course on self-interacting processes in Oxford in 2005 and 2007, and Nathanael Enriquez and Christophe Sabot for their invitation to lecture a minicourse on reinforced walks in Aussois in June 2009, both of which initiated these notes.

## References

- [1] M. Benaïm and P. Tarrès. Dynamics of vertex-reinforced random walks. *To appear in Annals of Probability*, 2011.
- [2] A. Bienvenue. Contribution à l'étude des marches aléatoires avec mémoire. *Doctoral dissertation, Université Claude Bernard-Lyon 1*, 1999.
- [3] D. Coppersmith and P. Diaconis. Random walks with reinforcement. *Unpublished manuscript*, 1986.
- [4] B. Davis. Reinforced random walk. *Probab. Theory Relat. Fields*, 84(2):203–229, 1990.
- [5] David A. Freedman. Bernard Friedman's urn. *Ann. Math. Statist.*, 36:956–970, 1965.
- [6] Bernard Friedman. A simple urn model. *Comm. Pure Appl. Math.*, 2:59–70, 1949.
- [7] David G. Kendall. Branching processes since 1873. *J. London Math. Soc.*, 41:385–406. (1 plate), 1966.
- [8] V. Limic. Attracting edge property for a class of reinforced random walks. *Annals of Probability*, 31:1615–1654, 2003.
- [9] V. Limic and P. Tarrès. Attracting edge and edge reinforced walks. *Annals of Probability*, 35, No. 5:1783–1806, 2007.
- [10] V. Limic and P. Tarrès. What is the difference between a square and a triangle? In *In and out of equilibrium 2. Series: Progress in Probability*, volume 60, pages 481–496. Birkhäuser, 2008.
- [11] R. Pemantle and S. Volkov. Vertex reinforced random walk on  $\mathbb{Z}$  has finite range. *Annals of probability*, 27:No.3,1368–1388, 1999.
- [12] T. Sellke. Reinforced random walks on the  $d$ -dimensional integer lattice. *Technical report 94-26, Purdue University*, 1994.
- [13] P. Tarrès. Vertex-reinforced random walk on  $\mathbb{Z}$  eventually gets stuck in five points. *Annals of Probability*, 32, No.3B:2650–2701, 2004.

- [14] Stanislav Volkov. Phase transition in vertex-reinforced random walks on  $\mathbb{Z}$  with non-linear reinforcement. *J. Theoret. Probab.*, 19(3):691–700, 2006.
- [15] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.